The role of miRNAs in mouse embryonic stem cells
A proteomics approach

Gyambibi-Barnett, Patricia

Awarding institution:
King's College London

The role of miRNAs in mouse embryonic stem cells: A proteomics approach

Patricia Gyambibi-Barnett

Supervisors

Professor Manuel Mayr, Professor Qingbo Xu, Dr Anna Zampetaki

Submitted for PhD

Table of Contents

1. Abbreviations 6

2. Figures 10

3. Tables 12

4. Abstract 13

5. Introduction 14

5.1. Cell potency 14

5.2. Unique properties of ESCs 16

5.3. Mouse embryonic stem cells 18

5.4. Transcription factors that regulate pluripotency 20

5.4.1. Oct4 20

5.4.2. Sox2 21

5.4.3. Nanog 22

5.5. Embryonic stem cell glucose metabolism 23

5.5.1. Glycolysis 24

5.5.2. Pyruvate kinase muscle isoform 2 (PKM2) 27

5.5.3. Mitochondria 30

5.5.4. Hypoxia 32

5.5.5. Embryonic stem cell cycle 33

5.6. Embryonic stem cell fatty acid metabolism 33

5.7. Embryonic stem cell amino acid metabolism 35

5.8. Non coding RNAs (ncRNAs) 36

5.8.1. MiRNA biogenesis 37

5.8.2. MRNAs in mESCs 39

5.8.3. miRNAs and mRNA translation 40

5.8.4. The function of miRNAs in ESCs 42

5.9. The microprocessor complex 45

5.9.1. DiGeorge critical region 8 (DGCR8) 45

5.9.2. Drosha 48

5.10. Argonaute 2 48

5.11. DGCR8$^{-/-}$ mESCs 50

5.12. MiRNAs and stem cell metabolism 51

5.13. Embryonic stem cell cycle (ESCC) regulating miRNAs 51

5.14. ESC Proteomics 54

5.14.1. Sample preparation 54
5.14.2. 2D Difference in gel electrophoresis (2D DIGE) 55
5.14.3. Protein identification 57
5.14.4. Ionization 57
5.14.5. Analysis and detection 58
5.15. Proton nuclear magnetic resonance (H^1 NMR) spectroscopy 59

6. Hypotheses **61**

7. Aim **62**
7.1. Analyses comparing WT and DGCR8 knockout mESCs 62
7.2. Analyses comparing DGCR8 null mESCs transfected with the individual ESCCs 62

8. Methods **63**
8.1. Cell culture 63
8.1.1. Derivation of MEF feeder layer 63
8.1.2. Mouse embryonic stem cells 64
8.1.3. Alkaline phosphatase staining 65
8.1.4. Cell transfection 66
8.2. RNA 67
8.2.1. RNA extraction 67
8.2.2. Reverse transcription PCR prior to conventional PCR 68
8.2.3. Polymerase chain reaction (PCR) 68
8.2.4. Reverse transcription prior to quantitative PCR for genes 69
8.2.5. Reverse transcription prior to quantitative PCR for miRNAs 69
8.2.6. Quantitative PCR (qPCR) using taqman assays 70
8.2.7. Quantitative PCR (qPCR) using SYBR green 70
8.3. Proteomic analyses 70
8.3.1. Cell protein extraction 70
8.3.2. Protein concentration measurement 71
8.3.3. 2D DIGE clean up 71
8.3.4. Silver staining 73
8.3.5. Analysis of 2-DE gels 74
8.3.6. In gel tryptic digestion 74
8.3.7. Mass spectrometry 75
8.3.8. Buffer exchange 76
8.3.9. Immunoblotting 76
8.4. Cell death (apoptosis) assay 77

8.5. Metabolomics 78
8.5.1. Perchloric acid (PA) extraction of cells 78
8.5.2. Proton nuclear magnetic resonance (H^1 NMR) 78
8.6. Seahorse bioscience 78
8.7. Statistical analysis 80
9. Results 81
9.1. WT and DGCR8$^{-/-}$ mESCs 81
9.1.1 Characterisation 81
9.1.2 DIGE 87
9.1.3 Validation of DIGE analysis (and associated proteins) of wild type and DGCR8$^{-/-}$ mESCs 91
9.1.4 Metabolic analysis 96
9.2 Transfection studies 99
9.2.1 Characterisation after transfection 99
9.2.2 Transfection DIGE 103
9.2.3 Validation of proteins from DIGE analysis and associated proteins in transfected cells 110
9.2.4 Metabolomics 116
10 Discussion 120
10.1 Characterization studies 121
10.1.1 DGCR8$^{-/-}$ mESCs may display increased differentiation 121
10.1.2 Sox2, but not Oct4 or Nanog display lower protein expression in the DGCR8$^{-/-}$ mESCs 121
10.1.3 DGCR8$^{-/-}$ mESCs display a slower proliferation rate 123
10.1.4 The overexpression of miR-291b in DGCR8$^{-/-}$ mESCs results in increased cell death 125
10.2 Proteins influenced by the ESCC miRNAs 126
10.3 Glucose metabolism 128
10.3.1 DGCR8$^{-/-}$ and WT mESCs display differential expression of glucose metabolism enzymes 128
10.3.2 Aldolase A's substrate FBP regulates PKM2 expression 131
10.3.3 The DGCR8$^{-/-}$ mESCs display a higher expression of TCA cycle enzymes 132
10.4 DGCR8$^{-/-}$ mESCs may use an alternative energy generation pathway to glycolysis 134
10.4.1 HIF1α expression in DGCR8$^{-/-}$ mESCs 137

10.4.2 Glutamic pyruvate tranaminase 2 (Gpt2) and Isocitrate dehydrogenase 1 (Idh1) are both indirectly regulated by the ESCC miRNAs 138
10.4.3 Threonine is secreted at higher levels by the transfected DGCR8$^{-/-}$ than those transfected with the control miRNAs 139
10.5 DGCR8$^{-/-}$ mESCs display higher expression of antioxidant proteins 140
10.6 Chaperone proteins 141
11 Limitations of the study and future work 143
11.1 Cell characterization 143
11.2 Technical difficulties 143
11.2 Future work 144
12 Conclusion 145
13 References 148

1. Abbreviations

2D DIGE	2D difference in gel electrophoresis
ADP	Adenosine diphosphate
Ago	Argonaute
Aldolase A/ALDOA	Fructose bisphosphate aldolase A
ANOVA	Analysis of variance
ATF6	Activating transcription factor 6
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CDK2	Cyclin dependent kinase 2
CDKN1A	Cyclin dependent kinase inhibitor 1
cDNA	Complementary DNA
CDS	Coding sequence
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1 propanesulfonate
CHOP	CCAAT-enhancer-binding protein homologous protein
CID	Collision-induced dissociation
CPT1	Carnitine palmitoyltransferase 1
Cy2/3/5	Cyanine dye 2/3/5
DGCR8	DiGeorge critical region 8
DHAP	Dihydroxyacetone phosphate
DMEM	Dulbecco's modified eagle's medium
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
ECAR	Extracellular acidification rate
ECCs	Embryonal carcinoma cells
ECD	Electron capture dissociation

EGCs	Embryonic germ cells
eIF4G	Eukaryotic initiation factor 4G
ELISA	Enzyme-linked immunoassay
ESC	Embryonic stem cell
ESCC	Embryonic stem cell cycle
ESI	Electrospray ionization
ETD	Electron transfer dissociation
FBS	Fetal bovine serum
FCCP	Carbonyl cyanide-*4*-(trifluoromethoxy)phenylhydrazone
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GLUT	Glucose transporter
GO	Gene ontology
Gpt2	Glutamic pyruvate transaminase 2
FAO	Fatty acid oxidation
FBP	Fructose-1,6-bisphosphate
^{1}H NMR	Proton nuclear magnetic resonance spectroscopy
HDAC1	Histone deacetylase1
HEK cells	Human embryonic kidney cells
hESCs	Human embryonic stem cells
HK2	Hexokinase 2
HIF1/2α	Hypoxia inducible factor 1/2α
HPLC	High performance liquid chromatography
IDH1/2	Isocitrate dehydrogenase 1/2
IEF	Isoelectric focusing
IL6	Interleukin 6
iPSCs	Induced pluripotent stem cells
LDHA	Lactate dehydrogenase A
LIF	Leukaemia inhibitory factor
LIFR	Leukaemia inhibitory factor receptor
lncRNA	Long non coding RNA
MALDI	Matrix assisted laser desorption/ionization
MDH2	Malate dehydrogenase 2

MEFs	Mouse embryonic fibroblasts
mESCs	Mouse embryonic stem cells
miRNA	MicroRNA
mRNA	Messenger RNA
Mw	Molecular mass
MS	Mass spectrometry
m/z	Mass/charge
NAD(H)	Nicotinamide adenine dinucleotide
ncRNAs	Non-coding RNAs
OCR	Oxygen consumption rate
Oct4	Octamer binding transcription factor 4
PA	Perchloric acid
PBS	Phosphate buffered saline
PCA	Principal component analysis
PCR	Polymerase chain reaction
PDI	Protein disulfide isomerase
PEP	Phosphoenolpyruvic acid
PERK	PKR-like endoplasmic reticulum kinase
PFA	Paraformalehyde
PFK1	Phosphofructokinase 1
PGAM1	Phosphoglycerate mutase 1
pI	Isoelectric point
PK	Pyruvate kinase
PKL	Pyruvate kinase liver isoform
PKM1/2	Pyruvate kinase Muscle isoform 1/2
PKR	Pyruvate kinase erythrocyte isoform
PMF	Peptide mass fingerprinting
Poly-A tail	Poly-adenylated tail
PPARα	Peroxisome proliferator-activated receptor alpha
PRDX3	Peroxiredoxin 3
Pre-miRNA	Precursor miRNA
Pri-miRNA	Primary miRNA
PRMT5	Protein arginine methyltransferase 5

qPCR	Quantitative PCR
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Ribonuclease
ROS	Reactive oxygen species
Rpm	Revolutions per minute
rRNA	Ribosomal RNA
RT-PCR	Reverse transcription PCR
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
siRNA	Small interfering RNA
snoRNA	Small nucleolar RNA
SOD1/2	Superoxide dismutase 1/2
TCA cycle	Tricarboxylic acid cycle
TDH	Threonine dehydrogenase
TPI	Triosephosphate isomerase
TRBP	Transactivating response RNA binding protein
tRNA	Transfer RNA
TXNIP	Thioredoxin-interacting protein 1
UCP2	Uncoupling protein 2
UPR	Unfolded protein response
UTR	Untranslated region
WT	Wild type

2. Figures

Figure 1: Embryonic stem cell cycle

Figure 2: Schematic representation of glucose metabolism in embryonic stem and differentiated cells

Figure 3: Maintenance of the pluripotent state in mESCs

Figure 4: Glycolytic pathway in ESCs

Figure 5: Switch between tetrameric and dimeric forms of PKM2

Figure 6: MicroRNA biogenesis

Figure 7: MicroRNA binding to 3'UTR of target sequence

Figure 8: Seed sequences of miRNAs that regulate the cell cycle in ESCs

Figure 9: Autoregulatory circuit between Drosha and DGCR8

Figure 10: Summary of 2D Difference in Gel Electrophoresis (2D DIGE) method

Figure 11: Summary of Mass Spectrometry on a 2D-DIGE gel

Figure 12: Summary of the proteomics workflow

Figure 13: Alkaline phosphatase staining

Figure 14: Expression of pluripotency markers

Figure 15: Genotype of wild type and DGCR8$^{-/-}$ mESCs and expression levels of miRNAs involved in the embryonic stem cell cycle

Figure 16: mRNA expression levels of cell cycle inhibitors

Figure 17: Cell numbers

Figure 18: Cell death detection assay

Figure 19: Representative DIGE gel image

Figure 20: Pathway analysis

Figure 21: Protein expression profile

Figure 22: Densitometry

Figure 23: Quantitative PCR profile

Figure 24: Quantitative PCR profile (2)

Figure 25: Seahorse analysis of wild type and DGCR8$^{-/-}$ mESCs

Figure 26: Metabolites in cell culture supernatants

Figure 27: Expression of pluripotency markers and cell cycle inhibitor in DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs

Figure 28a: Representative photos taken of transfected DGCR8-/- mESCs
Figure 28b: Average cell numbers of transfected DGCR8$^{-/-}$ mESCs
Figure 29: Representative DIGE gel image showing 58 differentially expressed proteins between DGCR8$^{-/-}$ mESCs transfected with miR-302d and the control miRNA
Figure 30: Principal Component Analysis (PCA)
Figure 31: Venn diagram of proteins from DIGE study
Figure 32: Pathway analysis
Figure 33: Protein expression of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs
Figure 34: Densitometry analysis of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs
Figure 35: Quantitative PCR profile of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs
Figure 36: Secretome analysis of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs
Figure 37: Intracellular concentrations of metabolites in DGCR8$^{-/-}$ mESCs transfected with a control miRNA, miR-294 and miR-302d
Figure 38: Gene expression of HIF1α
Figure 39: Role and mRNA expression of glutamic pyruvate transaminase 2 (gpt2) in DGCR8$^{-/-}$ and wild type mESCs
Figure 40: Simplified representation of glycolytic pathway in wild type mESCs
Figure 41: Glutaminolysis via the 'reverse TCA cycle' in rare cancer cells
Figure 42: Summary figure of possible glutaminolysis in DGCR8$^{-/-}$ mESCs

3. Tables

Table 1: Summary of denaturing agents in protein solubilising buffers

Table 2: Cell densities for transfections

Table 3: Mastermixes for transfections

Table 4: Forward and Reverse primers

Table 5: Taqman microRNA/gene expression assays

Table 6: Antibodies

Table 7: Proteins identified in WT vs DGCR8$^{-/-}$ mESCs DIGE analysis

Table 8: Average percentages (± standard deviation) of metabolites produced by WT and DGCR8$^{-/-}$ mESCs over 48 hours

Table 9: Proteins identified in DGCR8$^{-/-}$ mESCs transfected with ESCC miRNAs DIGE analysis

Table 10: Efficiency of transfections based on average Ct values from qPCR

Table 11: Average final cell numbers of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs

Table 12: Secretome analysis of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs

Table 13: Intracellular metabolite analysis of DGCR8-/- mESCs transfected with the ESCC miRNAs

Table 14: Proteins identified in both DIGE studies

4. Abstract

MicroRNAs are short non-coding RNAs that regulate gene expression at the post-transcriptional level. Studies in mouse embryonic stem cells (mESCs) where DGCR8; a major protein involved in their biogenesis has been knocked out, have shown that the global loss of canonical miRNAs results in cell cycle defects, delayed and reduced expression of markers of differentiation, and an inability to downregulate pluripotency markers upon differentiation. By conducting a 2D-DIGE study, comparing protein expression in wild type and DGCR8$^{-/-}$ mESCs, the aim was to study the effects that the loss of DGCR8 has on the proteome of mESCs when grown under proliferative conditions. The loss of DGCR8 in mESCs resulted in the deregulation of proteins with a chaperone function and those involved in glucose metabolism. Notably enzymes involved in glycolysis were reduced, whereas those involved in the TCA cycle were upregulated compared to wild type cells. mESCs are known to be highly glycolytic and the form of glucose metabolism used by cells has been linked to their capacity to differentiate. A second DIGE study was carried out on DGCR8$^{-/-}$ mESCs individually transfected with Embryonic Stem Cell Cycle specific (ESCC) miRNAs, to establish novel targets of these miRNAs and study their effect on the proteome. The study revealed that the ESCC miRNAs influence the expression of glucose metabolism proteins, notably Aldolase A, a key enzyme for glycolysis was identified in both studies as being an indirect target of the ESCC miRNAs. High resolution nuclear magnetic resonance spectroscopy further revealed differences in metabolism between the DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs and those transfected with a control miRNA, indicative of a switch from predominantly glucose metabolism in the wild type mESCs to glutaminolysis for energy generation in the DGCR8$^{-/-}$. Therefore the same miRNAs that control the embryonic stem cell cycle, also play a major role in the metabolic status of these cells, which may in turn play a role in the controlling the balance between pluripotency and differentiation. At the time of writing this is the first study using proteomic techniques to compare DGCR8$^{-/-}$ and wild type mESCs, and to explore the effect of the ESCC miRNAs.

5. Introduction

In 2000, when the human genome was successfully decoded it was discovered that approximately 98% did not code for proteins (1). One class of ncRNA that has been found to play an essential role for life are microRNAs (miRNAs), short non-coding RNAs that regulate gene expression at the post-transcriptional level. miRNAs have been found to influence the expression of approximately 60% of human protein coding genes (2) emphasizing their important role in gene regulation. The loss of miRNAs in a range of cell types, has revealed their diverse functions including regulating the cell cycle (3), cell proliferation (4) and cellular stress responses (5).

Embryonic stem cells (ESCs) have the unique property, of being able to differentiate into any embryonically derived cell type and can self renew, 'indefinitely' *in vitro.* Therefore, stem cell gene expression is tightly regulated to keep them eternally poised between pluripotency and differentiation, in response to external cues. Their unique phenotype has made them an important resource for the study of early development and for regenerative medicine. Thus, it is essential that the regulation of pluripotency and differentiation in stem cells is thoroughly understood. Studies have revealed that miRNAs play an essential role in controlling the stem cell phenotype.

The importance of miRNAs in stem cells has been revealed in studies where their biogenesis has been impaired, resulting in the global loss of miRNAs (6,7). Although miRNAs affect protein expression, many studies instead focus on the effects of miRNAs on mRNA expression (8).

5.1. Cell potency

Pluripotent stem cells are defined as 'cells with the potential to differentiate into any of the three germ layers; endoderm, mesoderm or ectoderm' (9). Three major types of pluripotent stem cells have been identified in fetal tissues;

embryonic stem cells (10), embryonal carcinoma cells (ECC) (11) and embryonic germ cells (EGC) (12) which share the properties of self renewal and the ability to differentiate into any embryonically derived cell, as well as the ability to form teratomas when injected *in vivo* (9).

ECCs are derived from teratocarcinomas, which imbues them with malignant properties (13), therefore although they provide an excellent model of early development and differentiation, they are unsuitable for clinical purposes. EGCs are derived from fetal gonads and eventually give rise to the gametes (14). This means that of the three major types of pluripotent cells that can be directly derived from *in vitro* sources, ESCs are the most suited for regenerative medicine.

Cells derived from earlier stages of blastocyst development than pluripotent cells, such as the 2-cell or 4-cell stage have the potential to develop into a total embryo/ whole organism and are termed totipotent (9). Cells derived from later stages of development, which have a more limited capacity for differentiation such as mesenchymal stem cells which can differentiate into osteoblasts or adipocytes amongst other cells, that can further develop into other cell types are termed multipotent.

These unique properties of pluripotent stem cells have made them an important resource for the development of regenerative medicine, as well as a simplified way of studying early development and disease progression.

By forcing the expression of certain pluripotency genes (15), miRNAs (16) or adding specific chemical cocktails to the differentiated cells (17) induced pluripotent stem cells (iPSCs) have also been created from mature cell types. iPSCs share the two unique properties of ESCs, of long term self renewal and pluripotency, however both human and mouse iPSCs do differ from ESCs in some respects. For example, particular iPSC lines display an 'epigenetic fingerprint', that is more similar to their cells of origin than ESCs and also display differentiation defects (18,19). It should be noted that certain differences

observed between ESCs and iPSCs have been attributed to variations in genetic background or viral gene insertions (19). By conducting experiments on mESCs and iPSCs with the same genetic background, the maternally imprinted gene cluster; *Dlk1-Dio3* was discovered to be repressed in the iPSCs, but not in the mESCs, this was also accompanied by differential expression of approximately 50 miRNAs, resulting in an inability of iPSCs injected into 4n blast cysts to develop into successful embryos (19).

5.2. Unique properties of ESCs

ESCs are derived from the inner cell mass of blastocysts and have two unique properties; they are capable of long term self-renewal though the longer they are kept in culture, the more likely they are to acquire mutations (20) and they can remain pluripotent *in vitro* under defined conditions. This unique phenotype is controlled by intrinsic factors regulated by extrinsic signals. The extrinsic signals primarily consisting of signals from neighbouring cells and the matrix that the cells are grown on, whereas the intrinsic factors include gene expression that is regulated at the transcriptional, translational, post-translational and epigenetic levels (21). Although hampered by ethical issues and the difficulties in controlling stem cell fate are yet to be completely overcome, embryonic stem cell research is one of the major forces driving the field of regenerative medicine forward and provides great insight into gene expression patterns and cellular behaviour both during proliferative and differentiation phases of cells.

The ability to self renew 'indefinitely', is in part due to the prolonged expression of the telomerase enzyme, which maintains the lengths of their telomeres over successive cell divisions (22). When differentiation occurs, the expression of telomerase and its associated enzymes rapidly declines, so that more mature cells have a limited number of cell divisions and eventually succumb to senescence (22).

Besides pluripotency and the ability to self renew, ESCs also have a very rapid proliferation rate, which is comparable to some cancer cell types (23) and

significantly faster compared to somatic cells (24). This is due to the ESC cell cycle; ESCs have a persistently high activity of the cell cycle regulators cyclin dependent protein kinase 2 (Cdk2) and cyclin A and E, which are known to drive the G1 to S phase transition (23) (Figure 1). Cdk2 has also been proposed to drive the G2 to M phase transition in a human cancer cell line and Xenopus cells (25), therefore its high expression may account for both of the very short G phases in ESCs. Unsurprisingly, this means that at any one time, a high proportion of ESCs within a population is in the S-phase or M phase (23). In stark contrast, the cyclin A and E proteins in differentiated cells are more tightly regulated and expressed at specific times during the cell cycle (23). One possible reason for this unusual cell cycle is to prevent an accumulation of differentiation signals, until they reach a threshold level that triggers the transition from self-renewal to differentiation (26,27).

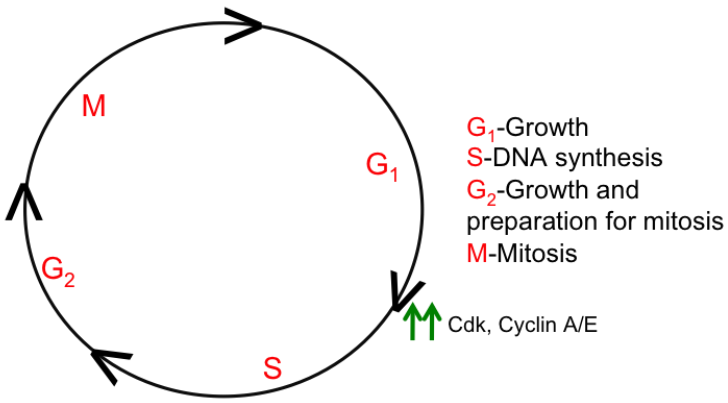


Figure 1 – Embryonic stem cell cycle. High expression of Cdk and Cyclin A/E promotes a rapid G_1/S phase transition.

In order to maintain this rapid rate of proliferation, ESCs mostly harness energy from aerobic glycolysis, as opposed to differentiated tissues, which derive the majority of their energy from oxidative phosphorylation or anaerobic glycolysis (28). Although, aerobic or anaerobic glycolysis results in the formation of less ATPs per molecule of glucose and a higher amount lactate, aerobic glycolysis provides energy much more promptly than oxidative phosphorylation as it bypasses the mitochondria (28) (Figure 2).

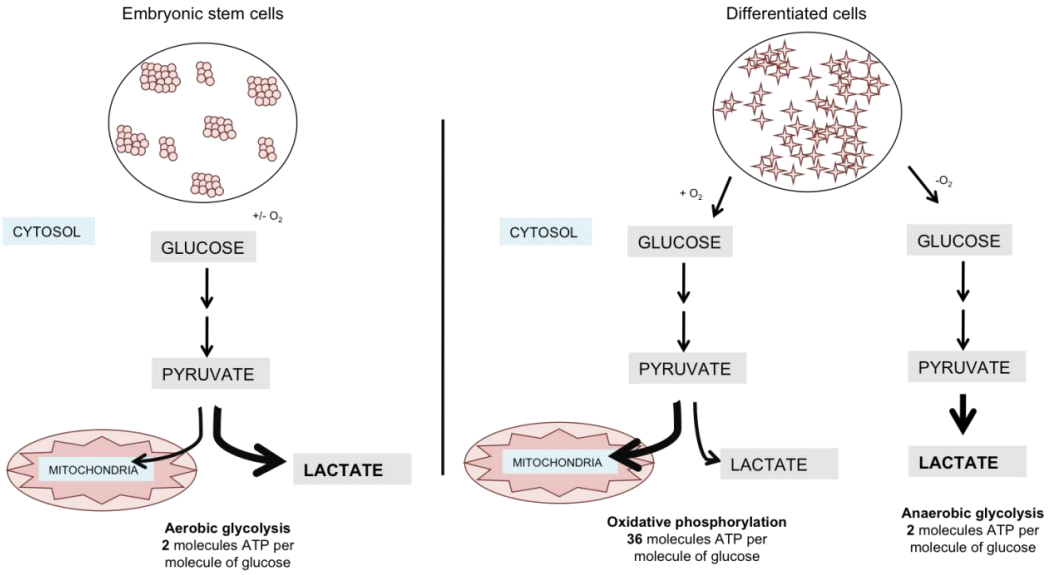


Figure 2 – Schematic representation of glucose metabolism in embryonic stem and differentiated cells. ESCs predominantly use the more rapid aerobic glycolysis in the cytosol for the generation of 2 molecules of ATP and a high concentration of the waste product lactate, as opposed to differentiated tissues which use either oxidative phosphorylation (mitochondria) or anaerobic glycolysis (cytosol), which generate 36 molecules and 2 molecules of ATP, respectively.

It has been suggested that as ESC proliferation is not limited by nutrients, under standard *in vitro/vivo* conditions, there has been no selective pressure to optimise for more efficient ATP production and that using glycolysis may have a protective effect on the cells, as a by-product of oxidative phosphorylation is increased ROS generation (28).

Recently it has been shown that the primary form of metabolism used by cells may control the balance between pluripotency and differentiation (29,30).

5.3. Mouse embryonic stem cells

Mouse ESC (mESC) lines were initially first propagated in the early 1980s, either on mitotically inactivated mouse embryonic fibroblast (MEF) feeder layers (31) or in the conditioned media of ECCs (32), which led researchers to believe that the MEF/ECCs were secreting a substance into the medium that was aiding the

mESCs in maintaining pluripotency. It was later discovered that this 'substance' was the cytokine myeloid Leukaemia Inhibitory Factor (LIF), a molecule which induces differentiation in M1 myeloid leukemic cells (33). The opposing effects of LIF on mESCs and M1 cells, is thought to be due to LIF stimulating different intracellular pathways in the two cell types (34).

LIF is an interleukin 6 (IL6) class cytokine that promotes the pluripotent state of mESCs by binding a receptor complex consisting of gp130 and the low affinity LIF receptor, resulting in the activation of the JAK/STAT3 pathway and the binding of STAT3 as well as other transcription factors such as Tcf3 and Klf2 to the enhancer region of the Oct4 gene (35), STAT3 activation also inhibits differentiation towards the mesoderm/endoderm lineage (36) (Figure 3). This is particularly important, as Oct4 along with Nanog and Sox2 have been classed as pluripotent genes. They are highly expressed in mESCs and upon their downregulation, differentiation ensues (37). Interestingly, studies have shown that the activation of STAT3 is critical for mESC pluripotency (34). LIF has also been shown to aid in the maintenance of the characteristic cell cycle of mESCs, by contributing to the rapid G phase of mESCs (27).

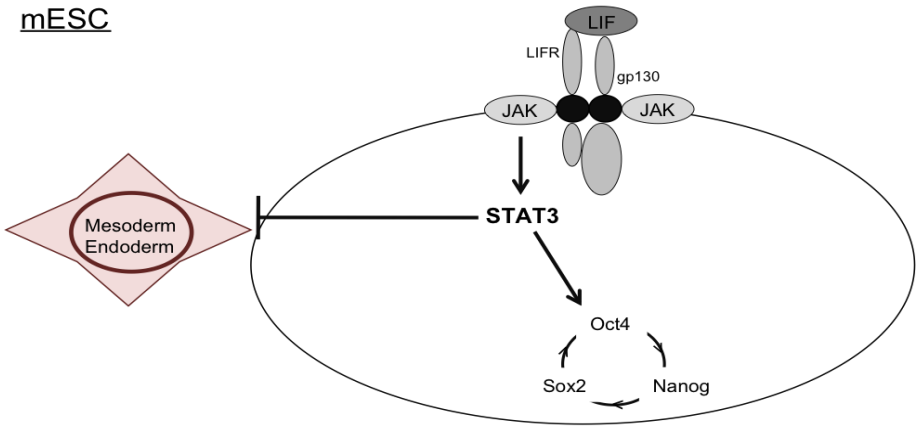


Figure 3 – Maintenance of the pluripotent state in mESCs. The binding of LIF to the LIF receptor (LIFR) results in the activation of the JAK/STAT3 pathway, subsequently the three pluripotency markers; Oct4, Nanog and Sox2 are activated, and differentiation is blocked. Adapted from: (36)

Although LIF is widely used in mESC culture to maintain an undifferentiated state, studies have shown that other members of the IL6 class of cytokines can also bind to the receptor complex and maintain pluripotency (34). However, as

LIF is highly expressed by the trophoectoderm *in vivo* and acts upon high affinity receptors in the inner cell mass (ICM) to maintain pluripotency, it best recapitulates the *in vivo* state. More recently, studies have found that LIF is dispensable for mESC pluripotency *in vitro* if it is replaced by a cocktail of enzyme inhibitors, which block pro-differentiative signalling such as the Map kinase pathway, indicating that the JAK/STAT3 pathway is not the only pro-pluripotency pathway in mESCs (38). Notably, when LIF is added to the cocktail of enzyme inhibitors, pluripotency is sustained for a longer period establishing its important role in the maintenance of pluripotency (38).

5.4. Transcription factors that regulate pluripotency

Oct4, Sox2 and Nanog are three highly expressed and conserved genes that imbue ESCs with the key property of pluripotency. Interestingly, in the developing blastocyst, to maintain the embryonic phenotype, pluripotent cells have to express Oct4 and Nanog, though not Sox2, which is thought, be due to presence of stable maternally derived Sox2 protein (38).

5.4.1. Oct4

Oct4 (octamer binding transcription factor 4) also known as POU5F1 is a highly conserved gene that is expressed during embryonic development in the epiblast and in ESCs, *in vitro*. The downregulation of Oct4 expression results in differentiation, therefore it is an important regulatory transcription factor involved in the maintenance of pluripotency (37). This is supported by the inability of Oct4 null embryos to outgrow in culture and their embryonic lethal phenotype (39), as well as the role of Oct4 (as one of the Yamanaka factors) in generating iPSCs (40,41). Forced Oct4 expression accompanied by a cocktail of chemical inhibitors in more mature cell types such as keratinocytes and fetal neural stem cells, has been found to be sufficient for the reprogramming process, albeit at a low efficiency, highlighting the role of Oct4 as a master controller of pluripotency (42).

Notably, of the four Yamanaka factors; Oct4, Sox2, Klf4 ad c-Myc, (43), Oct4 is the only factor which is absolutely crucial for iPSC formation, as the other factors can be substituted with homologous transcription factors (44). In order for mESCs to remain pluripotent, the level of Oct4 expression in mESCs is maintained within a threshold. Levels of expression either 50% below or above normal diploid levels result in differentiation either to the extra-embryonic endoderm and mesoderm lineages or trophoblast lineages, respectively (45). Notably, Oct4 over- or underexpression in ESCs directly represses the expression of transcription factors such as Rex1, which are associated with pluripotency, however when Oct4 is expressed within threshold levels, it activates Rex1 expression (46). Therefore Oct4 expression levels are tightly regulated in ESCs. This is at least partially controlled via an autoregulatory mechanism, however levels of expression are also controlled by the other 'pluripotency' genes, Nanog and Sox2 (47) (Figure 3) and other modulators such as miRNAs (48), in response to extrinsic signals.

5.4.2. Sox2

Similar to Oct4 null embryos, Sox2 null embryos are embryonic lethal (39) and knocking out Sox2 results in the loss of pluripotency in ESCs and differentiation towards the trophoectoderm lineage (49). Contrastingly, excessively high levels of Sox2 expression have been shown to promote differentiation towards the ectoderm and mesoderm lineages (50). This indicates that similar to Oct4, Sox2 levels are kept within a threshold in order to maintain the pluripotent state.

Studies have shown that Oct4 and Sox2 bind to enhancer sequences on each other (47) to form heterodimers in order to positively regulate and enhance each other's expression (51). They also act in concert to increase the expression of other pluripotency related genes including Nanog (49) (Figure 3). Surprisingly the expression of many Oct4/Sox2 target genes including Nanog are not affected in Sox2 null cells. This is thought to be due to a redundancy of Sox2 (49), which is supported by the fact that Sox2 can be substituted with Sox1 or Sox3 in the reprogramming of somatic cells (44). The major role of Sox2 in ESCs seems to be to maintain the high expression levels of Oct4, as exemplified in experiments

where Oct4 expression was forcibly kept at normal levels against a Sox2 null background resulting in maintenance of the pluripotent phenotype (49). The same effect was not observed when Nanog was overexpressed in the Sox2 null background cells, for unknown reasons, however it should be noted that Sox2 null cells maintain Nanog expression at similar levels to WT mESCs (49).

Although Sox2 is classed as a pluripotency gene it is also expressed by cells of the extra-embryonic ectoderm and precursor cells of the central nervous system, which divide rapidly (52,53).

5.4.3. Nanog

Nanog null embryos are embryonic lethal (54), nonetheless Nanog null mESCs have been generated, which surprisingly can remain pluripotent as long as the expression levels of Oct4 and Sox2 remain high. However, Nanog null mESCs are more susceptible to differentiation (51), reinforcing the importance of the expression of all the pluripotency factors for the maintenance of pluripotency.

In stark contrast to Oct4 and Sox2, overexpression of Nanog in the absence of LIF, results in the maintenance of the pluripotent state in ESCs, and causes them to be more resistant to retinoic-acid induced differentiation (37) in a dose dependent manner (51). This indicates that Nanog may play a dominant role in the maintenance of pluripotency, possibly through the direct downregulation of differentiation genes such as GATA4 and GATA6 (54) as well as via interactions with pluripotency promoting genes.

Similar to Oct4, Nanog can be used for one factor reprogramming when combined with chemical activators of sonic hedgehog signalling (involved in controlling early vertebrate embryonic development) (55).

Nanog's dominant role in maintaining pluripotency has led to the theory that it acts as a 'gatekeeper' for the onset of differentiation, as the downregulation of Nanog results in a concomitant downregulation of the other pluripotency genes Sox2 and Oct4, resulting in widespread differentiation (37). This theory is

supported by studies whereby Nanog RNAi has been used to forcibly repress Nanog expression in mESCs resulting in repression of Oct4 and Sox2 expression, conversely when Nanog was overexpressed, levels of Oct4 and other pluripotency associated genes, but not Sox2 increased (37). These results imply that Sox2 may be at the bottom of a 'hierarchy of pluripotency maintenance' whereby a drop in Nanog expression levels results in cells becoming less resistant to differentiation signals, leading to a decrease in Oct4 transcription levels (37) and a subsequent reduction in Sox2 levels (47,56). This concurs with more recent evidence that suggests that during iPSC formation, Sox2 is the first of the pluripotency genes to be switched on and in a hierarchal fashion, switches on other genes involved in the induction and maintenance of pluripotency (57).

Although, Nanog is proposed to play a major role in pluripotency, surprisingly it is not one of the core Yamanaka factors for reprogramming mESCs (though it is essential for gene based reprogramming of human ESCs) (15,41,58). However it has been discovered that unless murine iPSCs are expressing Nanog, they may not be completely reprogrammed, therefore Nanog is indispensable for the final stage in reprogramming (51), emphasizing its essential role in promoting pluripotency.

Although Oct4, Sox2 and Nanog are often cited as the master regulators of pluripotency, it should be noted that other genes such as Foxd3 (59), FGF4 and STAT3 (34) are vital for the maintenance of pluripotency, when mESCs are grown under normal conditions.

5.5. Embryonic stem cell glucose metabolism

Metabolism is directly or indirectly, intrinsically linked to every cellular process. Evidence has begun to emerge that the form of metabolism predominantly utilised by a cell plays a major role in its differentiation potential, and this knowledge has been applied to mESCs, which have an unusual form of glucose metabolism compared to their mature cell counterparts (Figure 2).

In contrast to mature cell types, which primarily depend upon mitochondrial oxidative phosphorylation to generate the energy required for cellular processes, stem cells use the faster, yet less efficient (in terms of the amount of ATP generated) aerobic glycolysis (28), which occurs in the cytosol (Figure 2). Although a definitive reason for this is yet to be uncovered, reasonable hypotheses have been formulated including the high availability of nutrients for stem cells resulting in there being no selective pressure for the cells to use oxidative phosphorylation, glycolytic intermediates feeding into anabolic pathways that are important for cell growth and glycolysis not generating ROS, which can be detrimental to cell survival (28). However more recently it has come to light that the form of glucose metabolism used by cells may play a role in controlling their pluripotent phenotype (60,61), this is supported by a switch from predominantly aerobic glycolysis to oxidative phosphorylation upon differentiation in ESCs (62), as well the switching on of glycolytic genes before pluripotency genes during iPSC generation (57,63). It has also been shown that glucose deprivation can result in delayed ESC differentiation, further demonstrating the importance of glucose as a substrate for normal the normal ESC phenotype (64).

5.5.1. Glycolysis

The GLUT family of transporters is comprised of thirteen members, which function to control cellular glucose homeostasis, by shuttling glucose into and out of cells (65). Interestingly, the expression of particular GLUT isoforms has been linked to the developmental status of both blastocysts and the ESCs derived from them (64). Of the GLUT isoforms, GLUT1 is essential for ESC viability indicative of its role as the main glucose transporter in ESCs, thus it is unsurprising that it is ubiquitously expressed during all stages of early development (64). The other GLUT isoforms are expressed sequentially over development, with both GLUT3 and GLUT8 showing higher expression during the early stages of differentiation and GLUT4 and GLUT2 displaying higher expression during the latter stages of development (64), which may allude to a role of the different glucose transporters in differentiation.

Glucose can be processed via two pathways for energy generation; glycolysis or oxidative phosphorylation (Figure 2). Both pathways are linked by the pyruvate kinase enzyme, which converts phosphoenolpyruvate into pyruvate, which can either be converted into lactate and two molecules of ATP in the cytosol during glycolysis, or be shuttled into the mitochondria (66) and converted into acetyl-CoA. Acetyl-CoA feeds into the Tricarboxylic Acid Cycle (TCA) cycle which coupled to oxidative phosphorylation, eventually results in the formation of 36 molecules of ATP, carbon dioxide and water (66) (Figure 4). Both the TCA cycle and glycolysis produce byproducts which can act as precursors for amino acid synthesis and nucleotide synthesis (66,67).

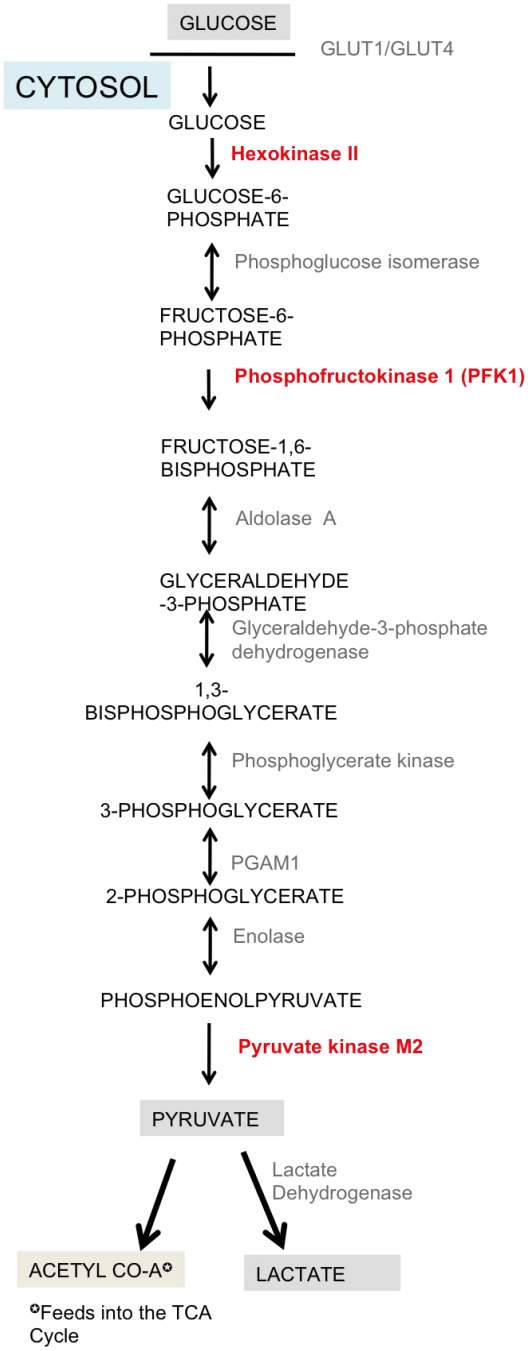


Figure 4 – Glycolytic pathway in ESCs. Enzymes in red catalyse the three rate limiting, irreversible steps of glycolysis Adapted from: (68)

Although both of the energy generating processes; glycolysis and the TCA cycle/oxidative phosphorylation can occur in the presence of oxygen, only the TCA cycle/oxidative phosphorylation actually requires oxygen in order to successfully generate energy (28) (Figure 2). The propensity of proliferating

cells including mESCs to use glycolysis instead of oxidative phosphorylation for energy generation, regardless of oxygen concentrations is termed the 'Warburg effect' (28,69).

Although mESCs are highly glycolytic, they are 'metabolically bivalent', and therefore have the ability to use both oxidative phosphorylation and glycolysis (70), this has been supported by studies which show that growing mESCs in high glucose concentrations results in not only increased glycolysis rates but also higher ROS production, possibly due to mitochondrial overload i.e. higher TCA cycle/oxidative phosphorylation activity (71).

Glycolysis has three rate limiting enzymes, which regulate the efficiency of the glycolytic pathway by catalysing the irreversible steps of glycolysis; Hexokinase 2 (HK2), Phosphofructokinase 1 (PFK1) and Pyruvate Kinase (PK). High expression of these enzymes is often indicative of active glycolysis. HK2, which commits glucose to the glycolytic pathway, can be positively regulated by insulin levels (72) and PFK1 by a number of factors including ATP and PEP (73), both of these enzymes predominantly play a major role in glycolysis (HK2 and PFK1). Whereas PK plays a major role in glycolysis but also in pathways directly associated with glycolysis which results in an upregulation of anabolic processes (74) (Figure 4).

5.5.2. Pyruvate kinase muscle isoform 2 (PKM2)

Four pyruvate kinase (PK) isoforms are present in mammals; the L and R isoforms which are expressed in liver and red blood cells, respectively, and the M1 and M2 isoforms (75), which are alternatively spliced forms of the muscle isoform, and have 96% similarity (76). The M2 isoform is expressed in all fetal tissues, but is progressively replaced by the M1, L and R isoforms in a tissue specific manner (77). Both the M1 and M2 isoform are expressed in highly proliferating tissues/cells, however the M2 isoform (PKM2) is predominantly present in highly proliferative embryonic tissues (Figure 4), whereas the M1 isoform (PKM1) is predominantly present in more mature proliferative tissues such as muscles and the brain (74). Whilst, PKL, PKR and PKM1 are able to form

stable, highly active tetramers (78), the one exon difference between PKM1 and PKM2 (79) results in PKM2's ability to intraconvert between a highly active tetramer and a low activity dimer. This intraconversion is under the control of binding partners such as amino acids (80) and substrates which are directly linked/involved in glucose metabolism such as Fructose 1, 6-bisphophate (FBP) which results in the formation of an active tetramer; (76) or ATP which results in the formation of a low activity dimer (81) (Figure 5). The other PK isoforms do not have binding sites, which allow for external regulation by binding partners, therefore they are constitutively expressed (82).

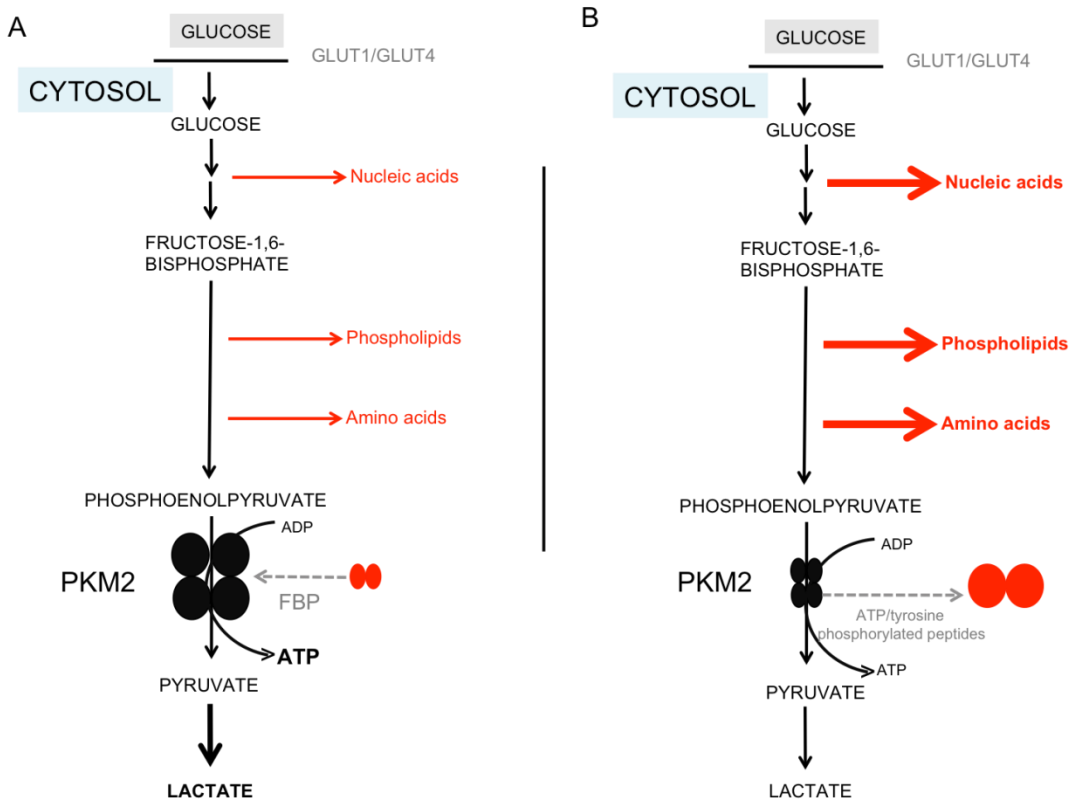


Figure 5 – Switch between tetrameric and dimeric forms of PKM2. A) PKM2 converts into a highly active tetrameric form upon the binding of Fructose 1, 6-bisphophate (FBP) resulting in increased glycolysis, B) PKM2 converts into a low activity dimer upon the binding of ATP or tyrosine phosphorylated resulting in increased channelling of substrates into anabolic processes. Adapted from: (83)

The switch between tetramer and dimer imbues proliferative tissues such as ESCs with a metabolic plasticity. In its tetrameric form, PKM2 has a high affinity for its substrate, phosphoenolpyruvate (PEP) resulting in the degradation of

pyruvate from glucose, leading to the production of lactate and two molecules of ATP (74), however in its dimeric form, pyruvate is instead channelled into anabolic pathways debranching from glycolysis, resulting in increased cell proliferation (74) (Figure 5). It is the tetramer/dimer ratio within cells, which determines the glycolytic fate of pyruvate (74,76). Although it has been established that the default state of PKM2 is as a tetramer, which allows for the rapid production of much needed ATP (84). It is interesting that in cancer cells the predominant form of PKM2 is as a dimer, which allows for the rapid growth of tumours (74) (at the time of writing, studies on the dimer, tetramer ratio in ESCs have not been conducted/published, though this is likely due to the dynamic nature of the ratio). It has been shown that when PKM2 is predominantly in its dimeric form, this is detected by the cell as 'nutritional stress' caused by low concentrations of glucose being converted to ATP, therefore glutaminolysis compensates for ATP production (74,84). It has also been suggested that upregulation of the dimeric form of PKM2, under low glucose conditions, aids in protection from apoptosis as shown by a downregulation of apoptosis markers (74).

Upon the receipt of signals, which inform the cell, that anabolic processes are required to further cell growth the intracellular balance between the dimeric and tetrameric forms of PKM2 is altered (74).

PKM2 has at least two binding sites, one of which allows FBP (activator) to bind as well as another, which allows tyrosine-phosphorylated peptides to bind (repressor) (82). In its default tetrameric state (84), PKM2 is bound by FBP. However upon the binding of tyrosine phosphorylated peptides, resulting in a conformational change of PKM2 to its dimeric form and the release of FBP (Figure 5), it is the concentration of FBP in the vicinity of PKM2 which determines whether PKM2 remains in its dimeric state or rebinds FBP, and is reactivated (82). This mechanism of regulation is thought to have evolved in fetal tissues, so that glucose is primarily used for cell growth i.e. anabolic processes, when cells are activated by the appropriate growth factor receptor protein tyrosine kinases (82).

Though embryonic stem cells are known to be highly glycolytic and express PKM2, it has only recently come to light that Oct4 and PKM2 are binding partners, with ectopic PKM2 expression directly enhancing the expression of Oct4 (85). Although it is still too early to postulate, what function this has in the cell and whether the two forms of PKM2 may have differing effects on Oct4 expression, it suggests a role for glucose metabolism in the regulation of Oct4 expression, and potentially pluripotency. PKM2 has also been shown to play a role in cell cycle regulation, as its loss results in the cell cycle stalling at the G1/S phase transition (86).

Notably, in studies on differentiated cells, when chemical inhibitors of glycolysis have been added, this has inhibited the reprogramming process and when glycolysis has been promoted, this has had the opposite effect (87,88).

5.5.3. Mitochondria

Mitochondria are double membrane organelles, of bacterial origin that play an essential role in energy, amino acid and fatty acid metabolism (89). By and large, ESCs do not predominantly use their mitochondria for energy production, instead relying on aerobic glycolysis, which occurs in the cytoplasm (28) and accounts for 80% of total cell glucose metabolism (30). However more recent studies have shown that although mESCs use aerobic glycolysis for energy generation, they are 'metabolically bivalent' and so have the ability to use either aerobic glycolysis or oxidative phosphorylation/TCA cycle to build precursors involved in cellular growth pathways and energy generation (70).

On the other hand, hESCs solely rely upon glycolysis for glucose metabolism; this has been demonstrated most effectively, in studies where both hESCs and mESCs have been grown in the presence of glycolytic inhibitors. hESCs begin to apoptose due to their inability to switch to oxidative phosphorylation/TCA cycle, whereas mESCs are able to compensate for their shortfall in energy production via glycolysis by using oxidative phosphorylation/ TCA cycle (70).

Although mESCs are able to use both glycolysis and the TCA cycle/oxidative phosphorylation, hESCs have elongated mitochondria, with distinct cristae compared to mESCs, which have more oval, irregularly shaped mitochondria indicative of the hESC mitochondria being more mature (70). Notably it has been shown that hESC mitochondria are functional and consume oxygen, however there is a greater uncoupling between glycolysis and oxidative phosphorylation in hESCs compared to mESCs (90). In stark contrast to differentiated mature cells, both hESCs and mESCs have underdeveloped and immature mitochondria (70). It has been discovered that once ESCs embark upon differentiation, they begin to upregulate genes associated with oxidative phosphorylation, the TCA cycle and ROS and reciprocally, downregulate the expression of glycolytic genes (63,70,91). Notably, this has not only been observed in ESCs, but also in multipotent cells such as mesenchymal and hematopoietic stem cells (90,92).

Although it is not well understood what controls the differences in mitochondrial metabolism between hESCs and mESCs, uncoupling protein 2 (UCP2), which acts a 'circuit breaker' between glycolysis and the TCA cycle/oxidative phosphorylation, has been implicated (61). During differentiation, UCP2 is repressed, resulting in increased oxidative phosphorylation (90). Reciprocally, high expression of UCP2 as observed in ESCs prevents the switch from a reliance on aerobic glycolysis to oxidative phosphorylation, and prevents the expression of early differentiation genes (90,93). However, the role of UCP2 in controlling the maintenance of pluripotency is still not completely certain (93).

Studies have shown that although mature cell types have significantly higher mitochondrial activity compared to ESCs, both cells have a similar mitochondrial mass (when normalised to cell mass) (90,94), and that the knockdown of a mitochondrial DNA polymerase (89) as well as the knockdown of the mitochondrial protein Growth factor erv1-like (Gfer) whose function is not completely understood (95) result in lower expression of the major pluripotency markers, as well as increases in the expression of early differentiation markers. The importance of mitochondrial function in regulating the balance between

pluripotency and differentiation has also been displayed in mESCs with deletion of the Cited2 gene; a transcriptional modulator, which have impaired mitochondrial function and increased glycolytic activity, resulting in defective differentiation and increased expression of the major pluripotency markers (30). Interestingly, Cited2 has been shown to directly affect the expression of Hexokinase 1 (30); an ubiquitous isozyme of the regulated Hexokinase 2 enzyme.

5.5.4. Hypoxia

Although *in vivo* embryonic cells grow under hypoxic conditions, standard *in vitro* conditions are at atmospheric oxygen tensions, which would indicate that oxygen tensions have no effect on pluripotency, however there are studies that suggest an optimal oxygen tension is required for proper mESC behaviour (96).

This is particularly interesting as mESCs contain immature mitochondria, with underdeveloped cristae in comparison to more mature cell types and hESCs, which contain more mature mitochondria (90). This may at least partially explain contradicting studies which show that hypoxia can either promote pluripotency, by activating the expression of hypoxia inducible factor 2α (HIF2α) which binds to and positively regulates Oct4 (96) or promote differentiation via the activation and binding of hypoxia inducible factor 1α (HIF1α) binding to and negatively regulating the LIF receptor (LIFR), preventing STAT3 activation and repressing pluripotency (97,98). Notably the opposing effects of oxygen on differentiation and pluripotency and the HIF proteins, is mirrored in other stem cell types, such as mesenchymal stem cells (99), pancreatic beta-cells (100) and hematopoietic stem cells (60).

Under hypoxic conditions HIF1α also activates the transcription of glycolytic genes, such as fructose bisphosphate aldolase A (Aldolase A) and phosphoglycerate mutase 1 (PGAM1) (101,102) (Figure 4) which has been linked to the maintenance of the pluripotent state. Hypoxia has also been implicated in increasing the efficiency of iPSC formation (103). Therefore it is clear that oxygen concentrations play a major, but as yet, not a completely

understood role in the balance between pluripotency and differentiation in mESCs.

5.5.5. Embryonic stem cell cycle

ESCs have a rapid cell cycle compared to their more differentiated counterparts, due to their very short G phases and reciprocally longer S phases. This not only has a profound effect on the pace of the cell cycle, but also results in ESCs ability to undergo at least two independent rounds of cell division in the absence of nutrients (104), and in stark contrast to differentiated cell types, not go into quiescence due to contact inhibition (105). Conversely, differentiated cells often arrest in G_0/G_1 phase upon nutrient starvation (105).

Unsurprisingly, in all cell types nutrient availability and both the pace and efficiency of the cell cycle are tightly intertwined. Growing mESCs in the presence of high glucose leads to the activation of key metabolic signal cascades that contain proteins such as PI3-K and Akt, which have wide-reaching roles in regulating glucose metabolism, cell proliferation and differentiation. This in turn activates Cdk complexes which enhance cell cycle function, resulting in increased cell proliferation (71). There is also evidence that ATP; one of the most vital end products of glucose metabolism, enhances the expression of the Cdk complexes, increasing mESC proliferation (106).

5.6. Embryonic stem cell fatty acid metabolism

Due to the strong links between the forms of metabolism predominantly used by a cell and the balance between pluripotency and differentiation, studies on the metabolome of cells have become more common. In the case of mESCs, a large scale unbiased study on two mESC lines and their differentiated progeny, showed that mESCs are characterised by an abundance of highly unsaturated fatty acids including secondary lipid messengers and inflammatory mediators, such as arachidonic acid and eicosapentaenoic acid, compared to their differentiated counterparts, and that the levels of unsaturated fatty acids

decrease upon differentiation (107). Conversely, the mature cell types had increased levels of saturated fatty acids, which have been implicated in the transport of fatty acids to the mitochondria for fatty acid oxidation (107). This correlates with numerous studies that show that upon differentiation, embryonic cell types increase their mitochondrial function (70). When the production of eicosanoids; a major group of unsaturated fatty acids involved in many processes including immunity and cell growth, was blocked, mESCs were unable to maintain pluripotency under pro-differentiation conditions (107) highlighting the importance of fatty acid metabolism for cell pluripotency. The regulation of the pluripotent state by fatty acids, has also been shown in hESCs (107).

Interestingly, linoleic acid, an unsaturated fatty acid that can be a precursor for eicosanoid production, has been shown to increase cell growth of mESCs by increasing the expression levels of the Cdk complexes and maintaining the rapid cell cycle (108).

Carnitine palmitoyltransferase 1 (Cpt1), a mitochondrial enzyme involved in fatty acid metabolism, has also been shown to play a major role in fatty acid metabolism in mESCs (109). In addition to impaired fatty acid metabolism, Cpt1 knockout mESCs have reduced ATP production, abnormally shaped mitochondria, impaired mitochondrial function, and succumb to apoptosis under hypoxic or low glucose conditions, further confirming the link between fatty acid metabolism and the normal mESC phenotype (109).

Notably UCP2, which is known to play a role in linking glycolysis to oxidative phosphorylation (61), has also been implicated in the mitochondria's 'choice' of either fatty acids or glutamine as a substrate for ATP generation and has been shown to be responsible for the metabolic switch from glucose to fatty acid metabolism in hypothalamic neurons (110).

5.7. Embryonic stem cell amino acid metabolism

Amino acids, known as the 'building blocks' of cells, play a vital role in the proliferation of all cells. Threonine in particular has been identified as an essential amino acid for mESC function (111) as out of the twenty amino acids, threonine is the only amino acid that is vital for mESC survival (111). mESCs also have high expression and activity of mitochondrial threonine dehydrogenase (TDH), which catalyses the first rate limiting step of threonine catabolism to glycine and acetyl-coA. Glycine can be used as a precursor for nucleotide synthesis, therefore in the absence of threonine dehydrogenase, there is a substantial decrease in DNA synthesis (111). Acetyl-coA can feed into the TCA cycle providing an alternative energy source (111). The study also showed that upon differentiation, there was a consistent increase in the concentration of guanosine, adenosine, inosine and threonine, in mature cell types (111). In stark contrast to mESCs, hESCs are dependent on methionine for cell survival (112). This is due to hESCs expressing a TDH pseudogene making them unable to catabolise threonine using the same pathway as mESCs (113).

The importance of threonine metabolism in mESC pluripotency has also been demonstrated in reprogramming studies, which have shown that the knockdown of TDH in cells being reprogrammed using the Yamanaka factors (41,44), lowers the efficiency of reprogramming (114). The opposite effect is observed if TDH activity is induced, and the reprogramming efficiency is further enhanced if PRMT5, a methyltransferase that is known to regulate TDH is also used in the reprogramming process (114).

Glutamine, which can be used as an alternative energy source by ESCs in the absence of glucose, has been linked to the rapid cell cycle in human cancer cells possibly providing evidence for links between glutamine metabolism and the rapid cell cycle, at least in mESCs (115).

In order to fully understand the mechanisms underlying the pluripotent state, it is imperative that there is a greater understanding of not only how genes

interact with each other, but also how different forms of metabolism affect the cell and how these different forms of metabolism interact and influence each other. Interestingly, both metabolism and pluripotency have been shown to regulated by miRNAs (48,116).

5.8. Non coding RNAs (ncRNAs)

In order for cells to express the correct genes both temporally and spatially in response to environmental or developmental cues, gene expression is tightly regulated. From transcription in the nucleus, to mRNA translation in the cytoplasm of cells and finally the action of the protein product; the 'message' that is being passed on, has to get past multiple checkpoints, can be heavily modified to repress or amplify the signal, and can be acted upon by small non-coding RNAs (ncRNAs) either activating or repressing translation (117).

Regulation of gene expression using small ncRNAs is a highly conserved process, which is involved, in various regulatory and signaling pathways. ncRNAs have been observed in plants as a defence mechanism against viruses (118), and in *C. elegans* they regulate embryonic development and lifespan (119). ncRNAs bind to their targets in a sequence specific manner resulting in translational repression. This can occur through a number of mechanisms including the binding of ncRNAs triggering cleavage of the target sequence, ncRNAs destabilizing the target sequence, most often by deadenylation (120), ncRNAs acting to prevent ribosome binding or through the ncRNA associated proteins such as the Argonautes (121). In the last decade, one category of small (~22nucleotides (nts)) ncRNAs called microRNAs (miRNAs) have been discovered to play a vital role in mammalian systems. This has been demonstrated most convincingly in studies where the miRNA biogenesis pathway has been impeded resulting in embryonic lethality in mice (6,122).

Although the first miRNA lin4, was discovered in 1993 in *C. elegans* and was found to negatively regulate the Lin14 protein by binding to the 3' UTR (untranslated region) of the *lin14* mRNA (123), it was not until 2000, when the

highly conserved let-7 miRNA was discovered in mammals that the field began to rapidly expand (124). So far, thousands of miRNAS have been discovered (119) and bioinformatic predictions suggest that mammalian miRNAs regulate at least 60% of all protein coding genes (2). Dysregulation of specific miRNAs has also been implicated in many disease states; examples include miR-15 and miR-16 in chronic lymphocytic leukemia (125) and miR-140 in arthritis (126).

5.8.1. MiRNA biogenesis

MiRNA genes can be found in clusters or as single genes, either in intergenic or intronic regions of the genome (127) and are transcribed in the nucleus by RNA polymerase II or III (21).

MiRNAs derived from intergenic regions of the genome are transcribed into primary transcripts with hairpin structures called primary miRNA (pri-miRNA) (21). Pri-miRNAs can contain several hairpin stem loops (128), and similar to mRNAs, these primary transcripts have 5' caps and 3' poly A tails. These pri-miRNA, are further processed into individual hairpins called precursor miRNA (pre-miRNA) (~ 70 nt) by the microprocessor complex. The microprocessor complex is made up of the Di-George syndrome critical region gene 8 (DGCR8) RNA binding protein (known as Pasha in Drosophila) in conjunction with Drosha, which provides a 'catalytic centre for the cleavage' (129) of the 5' cap and 3' poly-A tail, as well as separating multiple stem loops. The pre-miRNA is then transported from the nucleus to the cytoplasm, by exportin-5 and Ran-GTP, where the RNAse II enzyme, Dicer removes the hairpin loop and cleaves the pre-miRNA into a ~21 bp duplex with 2nt overhangs on its 3' ends (Figure 6). Of the two mature strands, one is known as the guide strand and is loaded onto the RNA-induced silencing complex (RISC) by Dicer (21). This is usually the strand with the weakest base pairing at its 5' end (130) and it targets specific mRNAs in a sequence dependent manner. In most cases the other strand, passenger/star strand has no function and is degraded (21). However, 80% of human miRNAs have highly conserved star strands, and a smaller percentage have higher or similar concentrations as the guide strand, indicating that the star strand may be

functional, and have targets of its own (131), such as miR-233* in mammalian myeloid progenitor cells which targets a number of mRNAs in the insulin-like growth factor 1 receptor/phosphatidylinositol 3-kinase pathway (132). Both Dicer and Exportin 5, have been suggested as rate limiting steps, for miRNA function (133).

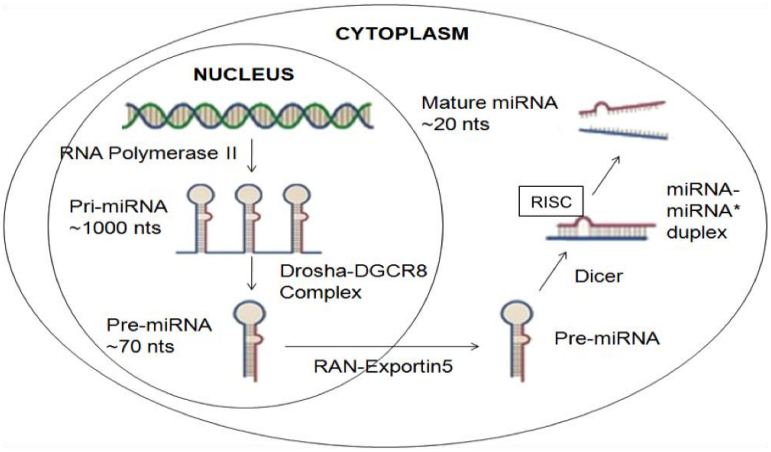


Figure 6 – miRNA biogenesis. Intergenic miRNAs are transcribed in the nucleus by RNA polymerase II into primary miRNA (pri-miRNA), which undergoes two sequential I steps by the Drosha-DGCR8/Microprocessor complex in the nucleus and Dicer in the cytoplasm, to form a miRNA duplex, which is recognised and cleaved by the RISC.

Studies have shown that the majority of miRNAs (~80%) are derived from intronic regions of the genome and are transcribed using the endogenous promoter of the gene transcription unit that the intron is located in (134). If the intron is within a protein coding gene, then often the intronic miRNA is found in an antisense orientation indicating that it is its own independent transcription unit, in some cases the intronic miRNA is found in the sense orientation and is co-transcribed with the protein coding gene (135). Some of the host genes that contain intronic miRNA are non-protein coding, suggesting that they exist solely for miRNA production (136).

The biogenesis of miRNAs derived from introns is very similar to miRNAs derived from intergenic regions of the genome, however due to the nature of introns it also involves the spliceosome (136). Originally it was thought that miRNA sequences were spliced by the spliceosome before the microprocessor

complex cleaved the pri-miRNA (137). However, there are studies, which suggest that intronic miRNAs can be processed from unspliced/partially spliced introns by the microprocessor complex, before splicing occurs, without any damage to the protein/mature mRNA sequence (136,138). Intronic miRNA biogenesis pathways that are not dependent on the microprocessor complex or its component parts and so bypass pri-miRNA production have been identified in *C.elegans* and *Drosophila* (139). More recently mammalian miRNAs that are produced via non canonical pathways, which are not only microprocessor independent, but also splicing independent, therefore only requiring Drosha for their biogenesis have also been discovered (140).

Recently, other DGCR8 independent pathways for miRNA biogenesis have been identified, that transcribe functional miRNAs from other sources including transfer RNAs (tRNAs) (141-143).

5.8.2. MicroRNAs in mESCs

Studies on mouse embryonic stem cells (mESCs) that no longer express the majority of their miRNAs due to the deletion of a protein (DGCR8) involved in their biogenesis (Figure 6), have an altered cell cycle, where cells accumulate in the G_1/S phase of the cell cycle, a slow proliferation rate compared to WT cells and an inability to downregulate genetic markers of pluripotency (6). This indicates that miRNAs play a vital role in the mESC cycle and in the differentiation status of the cell.

By re-expressing miRNAs in DGCR8 knockout mESCs, 14 miRNAs have been identified that can rescue the cell cycle phenotype. However, the miRNAs involved in the major regulation of mESC differentiation are yet to be identified (6). Large-scale transcriptomic studies have been conducted on both WT and DGCR8$^{-/-}$ mESCs and on DGCR8$^{-/-}$ cells transfected with the ESCC cluster of miRNAs amongst others, to discern potential targets of the ESCC miRNAs (144). GO analysis has revealed that the top pathways affected by these miRNAs are unsurprisingly involved in the cell cycle, epigenetic mechanisms, apoptosis and

immune responses (145-149). However, as the major read out of miRNA function is at the protein level, this provides an incomplete picture of the targets of these miRNAs, and their role in the ESC phenotype. This has been demonstrated most effectively in a recent study using an unbiased, large-scale approach to study the targets of *C.elegans* miRNAs, which revealed that the majority of miRNA interactions identified, did not have perfect seed sequence complementarity, and therefore would have not been easily identified using traditional bioinformatic, prediction methods (150).

5.8.3. miRNAs and mRNA translation

miRNAs are mediators of genetic repression and in rare cases, they can indirectly cause genetic activation during the cell cycle (117). They function at the post-transcriptional level by binding in a sequence specific manner to mRNA. Their importance in gene expression has been revealed in studies where their biogenesis has been disrupted, resulting in early embryonic lethality in murine models and abnormal differentiation patterns in cellular models (6).

Although the results of miRNA binding to their target sequences are known, the actual process of gene regulation by miRNAs is not completely understood due to varying experimental results showing that both the elongation and initiation stages of translation can be affected (151). However certain aspects of mRNA silencing are generally agreed upon such as the formation of the RISC being coupled to pre-miRNA processing by Dicer (152).

The mature strand guides the RISC complex to the target sequences (6), whereby if the guide strand and the target sequence are completely complementary then in mammals, Argonaute 2 (Ago2), an RNAse enzyme will cleave the target sequence between bases 10 and 11 so it cannot be translated (153) (Figure 7). However in animals, the guide strand and target sequence are often not perfectly complementary (131), consequently the mRNA sequence is not destroyed rather it is repressed (152). In the case of imperfect complementarity, the miRNA and its target sequence have to base pair in specific configurations in order for

effective repression to take place, in many cases the miRNA needs to perfectly base pair to its target sequence at the miRNAs 5' end, this area of perfect complementarity is known as the seed region and usually spans nucleotides 2-8 (154). If the seed region binding is not perfect, this can be compensated by strong binding between the 3' ends of the miRNA and target sequence (154) (Figure 7).

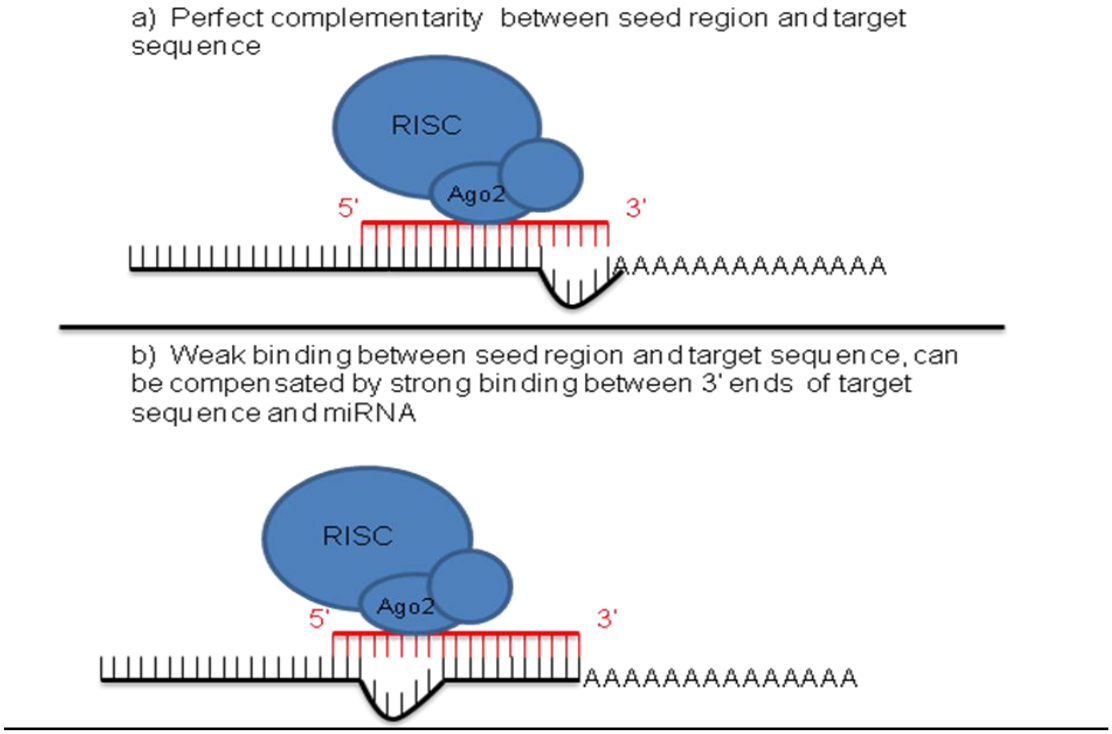


Figure 7 – miRNA binding to 3'UTR of target sequence. A) Perfect complementarity at seed region between miRNA and 3' UTR of target sequence results in cleaving of target, B) Weak binding at seed region, can be compensated by strong binding between 3' ends of miRNA and target sequence and results in target repression

Many miRNAs are required per target sequence for effective repression (152), once translation has been repressed, the target sequence and RISC are stored in P-bodies via an unknown mechanism (151,155).

miRNAs have been shown to repress translation in number of ways, however there is still debate over which method, if any, is the dominant method of translational repression and whether these methods occur simultaneously or individually (151). The methods include miRNAs interacting with the 40S and

60S subunits of ribosomes either preventing them from binding or causing them to drop off the target mRNA(131) or, RISCs competing with essential 5' cap binding proteins such as eIF4G preventing translation (156).

As well as directly repressing translation, miRNAs have also been shown to affect target mRNA levels by triggering deadenylation of the target mRNA sequence, resulting in decapping, disrupted translation, and eventually leading to degradation of the target sequence (131,157,158).

Interestingly, some miRNAs have been found to reside and function in the nucleus, indicating that once they are cleaved by Dicer in the cytoplasm, they are transported back into the nucleus. For example miR-709 is highly expressed in the nucleus of a number of cells including adipose derived fibroblasts and Human Embryonic Kidney (HEK) cells, and represses the expression of miR-15a/16-1 by preventing its maturation (159) and miR-122, one of the most abundant miRNAs in the liver (160) not only targets mRNA in the cytoplasm, but also targets mRNAs in the nucleus (161).

5.8.4. The function of miRNAs in ESCs

Studies in mouse embryonic stem cells (mESCs), have revealed that when Dicer is knocked out, resulting in an almost complete loss of mature miRNAs, although the mESCs are viable, they have a lower proliferative capacity i.e. grow much slower than their WT counterparts (6). Cell cycle analysis has revealed that the Dicer null mESCs have an altered cell cycle, where there is an accumulation of cells in the G_1 and G_o phases and a reciprocal decrease in the number of cells reaching the G_2/M phase (162). Dicer null mESCs are able to form embryoid bodies (EBs) (6) but, they fail to express genetic markers of differentiation such as HNF4A, an early marker of endoderm and BMP4, an early mesodermal marker (7).

In reference to their cell cycle and ability to differentiate, DGCR8$^{-/-}$ mESCs have a similar though less severe phenotype compared to Dicer knockout mESCs.

DGCR8$^{-/-}$ mESCs also fail to silence the pluripotency genes such as Oct4, Sox2 and Nanog (6). This was especially apparent when differentiation was forced using retinoic acid, which resulted in the expression of markers of differentiation in DGCR8 null mESCs as well as prolonged expression of the pluripotency genes. In contrast, retinoic acid was unable to force the expression of differentiation markers in Dicer-null mESCs (6,162). The proliferation defects in Dicer null mESCs can be rescued over time which has been attributed to some of the cells accruing mutations/stable genetic changes that can overcome the proliferation defects and outgrow the other cells, contrastingly DGCR8 null mESCs cannot overcome this proliferation defects for unknown reasons (163).

By reintroducing non-canonical miRNAs into Dicer$^{-/-}$ mESCs; two non canonical miRNAs, miR-320 and miR-702 that promote the G_1 to S phase transition have been discovered (8,164). This partially accounts for the differences in severity of phenotype between the DGCR8 and Dicer knockout mESCs. It is noteworthy that the expression of these two miRNAs and other as yet unidentified non-canonical miRNAs, is insufficient to completely rescue the defects in DGCR8$^{-/-}$ mESCs, implying that canonical miRNAs play a greater role in maintaining the mESC phenotype than non-canonical miRNAs.

These results indicate that miRNAs play a vital role in the differentiation and cell cycle of mESCs. In contrast to DGCR8, Dicer is involved in both the siRNA and miRNA biogenesis pathways. Thus the miRNA pathway plays a similarly predominant role in the phenotypes of the Dicer and DGCR8 null mESCs, whereas siRNAs may account for the difference in phenotypes between Dicer null and DGCR8 mutants(163).

The ability of specific miRNAs to restore the cell cycle defects in DGCR8 knockout cells (26) suggests that there are specific miRNAs in ESCs that promote the cell cycle and particularly play a vital role in the transition from the G_1 phase (26).

The phenotype of DGCR8$^{-/-}$ hESCs is yet to be uncovered; nevertheless it has been predicted to be similar to that of mESCs. Evidence for this has been provided by

hESCs and mESCs having similar miRNA profiles. For example, a cluster of several mESC specific miRNAs known as the miR-295 cluster; miR-290, miR-291a, miR-291b, miR-292, miR-293, miR-294 and miR-295, have orthologs in hESCs, the miR-371 cluster; miR-371, miR-372, mir-373 and miR-373* (165). Both clusters have the same seed sequence and share targets (122,166) (Figure 8). Interestingly in hESCs, the miR-371 cluster is in fact expressed at relatively low levels, instead another cluster of miRNAs, the 302 cluster; miR-302a, miR-302b, miR-302c and miR-302d, which also has the same seed sequence as the miR-295 and miR-371 clusters, is the most highly expressed cluster (Figure 8).

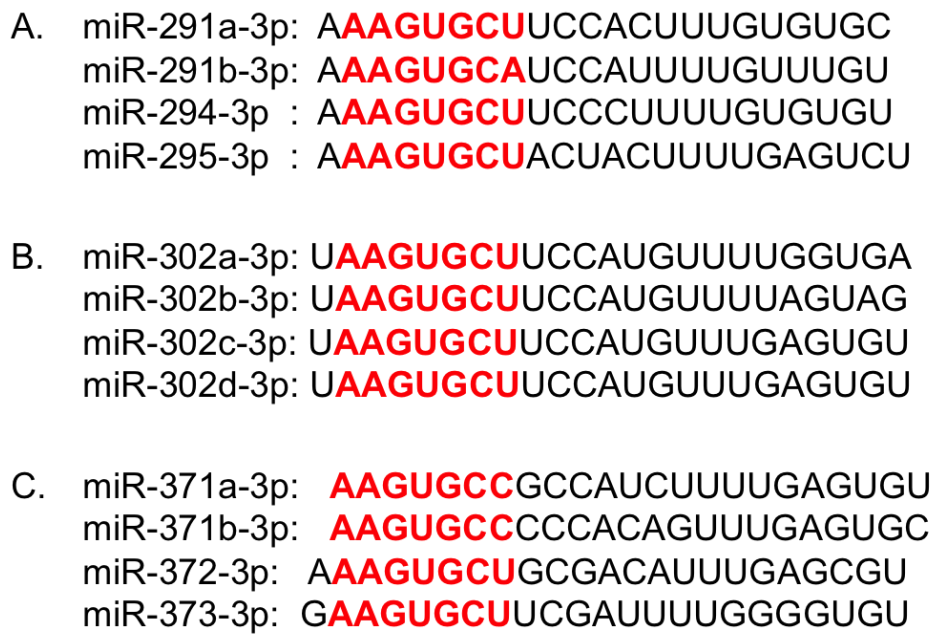


Figure 8 – Seed sequences of miRNAs that regulate the cell cycle in ESCs. A) Highly expressed mouse miR-295 cluster B) Highly expressed human miR-302 cluster c) Human miR-371 cluster

The transcription factors Oct4, Nanog and Sox2 act together and are central to the establishment and maintenance of 'stemness', as their down regulation initiates and marks the beginning of ESCs differentiation program. Therefore it is telling that they and one other transcription factor, Tcf3 control approximately 20% of all known miRNAs (48) in mESCs, either by occupying the promoters of active miRNAs that have a role in maintaining pluripotency or, by occupying the

promoters with other repressive complexes of miRNAs that are normally upregulated in differentiated cells, but silent in ESCs (48,166).

There is also a general increase in miRNA expression in hESCs and mESCs, following the loss of pluripotency and subsequent differentiation (166).

5.9. The microprocessor complex

5.9.1. DiGeorge critical region 8 (DGCR8)

DGCR8 was first identified in humans as a protein that is deleted in patients suffering with a rare disease called DiGeorge syndrome (167). The disease is caused by de novo or autosomal dominantly inherited deletions in the 22q11.2 chromosome region, as the genes deleted play a role in many biological systems, the symptoms of DiGeorge syndrome are diverse and can include cardiac abnormalities, hypocalcaemia and the loss of the thymus gland (167). Unlike the DGCR8 null mice which are early embryonic lethal, sufferers of DiGeorge syndrome are heterozygous for the deletion so produce a sufficient amount of DGCR8 for survival, this corresponds with studies conducted on mice which are haploinsufficient/heterozygous ($DGCR8^{+/-}$) for DGCR8 which are viable and develop neuronal defects postnatally (168). The neuronal defects observed in $DGCR8^{+/-}$ mice may explain some of the learning difficulties observed in some DiGeorge sufferers (169), however although DGCR8 maps to the deleted region in DiGeorge syndrome, it is yet to be confirmed whether the decrease in miRNAs is a direct cause of the disease symptoms(170). DGCR8 has also been implicated in other diseases such as in the inherited neurodegenerative disorder; fragile X-associated tremor/ataxia syndrome (FXTAS), in which the expanded CGG repeats in the fragile X mental retardation 1 (FMR1) mRNA bind to DGCR8, inhibiting the miRNA biogenesis pathway, resulting in neuronal cell dysfunction (171).

The main role of DGCR8 as a critical component of the microprocessor complex is to recognise the substrate pri-miRNA, in the nucleus, which is then cleaved by its partner Drosha (129). DGCR8 consists of two tandem RNA binding domains in the C terminus, which facilitate binding to Drosha and pri-miRNA, and a WW domain at the N terminus which plays a role in the nuclear localisation of DGCR8 (172). Although the DGCR8 protein can bind RNA (single stranded, double stranded and random hairpin transcripts) in a non-specific manner (173), it has recently been discovered that DGCR8 has the ability to form a homodimer, creating a functional RNA binding heme domain, which resides in the central region of DGCR8, to specifically bind to pri-miRNA (174). However, it is not fully clear whether the formation of the DGCR8 dimer is necessary for the successful formation of the microprocessor complex (174).

DGCR8 can also undergo post-translational modifications, in order to modulate its function, such as deacetylation by HDAC1 of lysine residues in its RNA binding domains, which increases its affinity for pri-miRNAs (175) and phosphorylation which stabilises the protein and in the human cells studied; Human Embryonic Kidney (HEK) 293 and HeLa, promotes a proliferative phenotype (176).

Further studies have suggested that as well as being directly involved in the miRNA biogenesis pathway via the Microprocessor complex, DGCR8 may also associate with unidentified endonucleases as part of the small nucleolar RNA (snoRNA) biogenesis pathway (177).

Unlike Drosha and Dicer (21,178) which have been reported to have a role in other small non coding RNA biogenesis pathways, DGCR8 is unique to the miRNA biogenesis pathway, therefore knocking it out allows for studies of the effects of the global down regulation of miRNAs (3). By reintroducing individual or clusters of miRNAs back into DGCR8 null cells, the functions of the miRNAs can be elucidated, without the interference of the other canonical miRNAs (3).

Recent studies have revealed that although DGCR8 and Dicer do not seem to play a direct role in regulating or creating long (>200 nucleotides) non coding RNA

(lncRNA), the loss of either protein, which results in the loss of the majority of miRNAs also, results in decreased expression of lncRNAs, although they lack 3' UTRs (177,179).

Within the nucleus, Drosha and DGCR8 regulate each other. Drosha suppresses DGCR8 expression in the nucleus post-transcriptionally, thereby indirectly regulating the production of mature miRNAs (180). Interestingly, as the 3' UTR of DGCR8 is hypothesised to have miRNA binding sites and DGCR8 mRNA have a hairpin structure, it is wholly possible that the post transcriptional down regulation of DGCR8 is mediated by miRNAs (180). The DGCR8 protein stabilises the Drosha protein, therefore as well as playing a functional role in miRNA biogenesis, DGCR8 also facilitates the successful formation of the microprocessor complex (Figure 9).

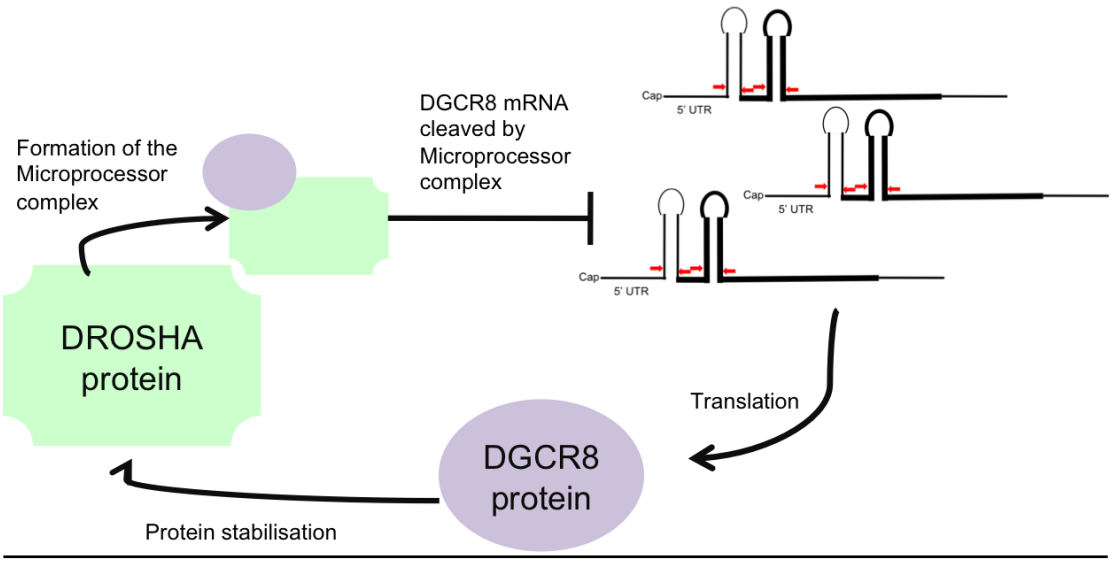


Figure 9 – Autoregulatory circuit between Drosha and DGCR8. The two components of the microprocessor complex regulate each other; the Drosha protein downregulates DGCR8 posttranscriptionally, and the DGCR8 protein stabilises the Drosha protein. Adapted from: (180)

Thus, Drosha and DGCR8 are involved in a complex auto regulatory circuit, whereby if DGCR8 levels are high, the microprocessor complex will post transcriptionally down regulate it, reducing the stability of Drosha, and lowering the activity of the microprocessor (180) (Figure 9). Conversely in a study in mESCs with reduced Drosha activity, there was an increase in DGCR8 expression

(138). This may be to allow for extra regulation of miRNA processing, preventing miRNAs from being expressed until the appropriate developmental signals are received (181).

The microprocessor complex is also regulated extrinsically by developmental or environmental cues, for example in reproductive cancer cell lines, cellular stress has resulted in p53 (tumour suppressor gene) interacting with Drosha to upregulate the production of specific miRNAs that suppress cell growth (182).

5.9.2. Drosha

Drosha is a class two, RNAse III endonuclease that resides in the nucleus and as an integral part of the microprocessor complex (along with DGCR8) aids in the maturation of miRNAs (183). Structurally Drosha has tandem RNase III domains, and a double stranded RNA binding domain. Akin to DGCR8, Drosha RNAse activity is unspecific, however unlike DGCR8, Drosha requires binding to DGCR8 to confer high specificity for the processing of pri-miRNAs (138,178). Therefore it has been suggested that Drosha may play a role in other ncRNA pathways. It has already been discovered that Drosha is integral to the DNA damage RNA (DDRNA) biogenesis pathway (184). These ncRNAs have the sequence of the damaged DNA locus and use this to signal the presence of DNA damage, so that repair can take place (185).

As an endonuclease, Drosha also plays other roles outside the miRNA biogenesis pathway such as the processing of pre-ribosomal RNA (186). Drosha has also been implicated in the processing of a murine, unspliced, polyadenylated, nuclear ncRNA known as meiotic recombination hot spot locus (*mrhl)* with a possible function in the nucleolus (187).

5.10. Argonaute 2

The Argonaute (Ago) protein family as part of the RISC complex, plays an integral role in RNA silencing processes, influencing both miRNA and siRNA

function (155). miRNAs and siRNAs guide Ago proteins as part of the RISC complex, to their target sequence where the ncRNAs bind via complementary binding, resulting in target degradation or translational repression mediated by Ago (188). The importance of Ago in mESCs is exemplified in studies where Ago has been knocked out, resulting in an inability of cells to synthesise miRNAs and siRNAs, and undergoing rapid apoptosis due in part to the upregulation of apoptotic genes due to the loss of ncRNAs (189).

It should be noted that Ago proteins have been found to interact with other ncRNAs, including transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) (1).

In humans and mice, four ubiquitously expressed Ago family members; Ago1-4, can function in the RISC complex interchangeably (189), though Ago2 is preferentially incorporated into the RISC complex for miRNA function (21). Although the members of the Ago family are highly similar, Ago1 and 2 are more effective at utilizing perfectly matched siRNAs than the other two family members, and there are a small percentage of miRNAs that are capable of being preferentially loaded onto particular Ago family members within RISC complexes (189,190). There is conflicting evidence that although all of the Ago proteins can bind to siRNA and miRNA, only Ago2 is able to cleave target mRNA sequences (191,192).

ncRNA interaction with Ago may be a rate limiting step for ncRNA function as ncRNAs can compete for Ago binding, mediating their function (133). Ago2 proteins are also regulated by a specific autophagic mechanism, which can degrade both the Ago2 and Dicer protein regulating miRNA loading into the RISC complex (193).

Though the proteins that make up the RISC complex, are yet to be identified (194), it has been discovered that the minimal requirements for the RISCs function as a gene silencer are the Ago protein and the mature miRNA (195). The guide strand is not only important for the identification of the target sequence, but its association with Ago also aids in the stability of the RISC (131).

As well as playing a major role in the RISC complex, Ago2 has also been shown to play another role in an alternative miRNA biogenesis pathway (127,188). In this pathway, after pre-miRNAs are exported from the nucleus to the cytoplasm, they are recognised by a preformed complex formed of Ago2, Dicer and the transactivating response RNA-binding protein (TRBP) (196), which nicks a section of the passenger strand, Dicer then cleaves the stem loop resulting in the formation of a miRNA/miRNA* duplex (127). At this point the Ago/Dicer/TRBP complex dissociates and it has been hypothesised that the free Ago2 is then free to be incorporated into the RISC complex (127). The nicking of the passenger strand of the pre-miRNA by the Ago/Dicer/TRBP complex is thought to aid in the both the identification of the guide/mature miRNA strand by the RISC complex (127).

5.11. DGCR8$^{-/-}$ mESCs

Of the individual components that are involved in the miRNA biogenesis pathway, DGCR8 is the only one, which is unique to this particular ncRNA biogenesis pathway; therefore by knocking it out, researchers have been able to study the effects of global miRNA loss (6). It should be noted that of course, in the absence of DGCR8, miRNAs generated via non-canonical pathways are still produced, however as these make up approximately 1% of the miRNAs produced by cells, that the DGCR8$^{-/-}$ mESCs are relatively miRNA free (197).

DGCR8$^{-/-}$ mESCs provide a valuable resource for studying the functions and targets of individuals or clusters of miRNAs, via the re-expression of these miRNAs in DGCR8$^{-/-}$ mESCs, for example transfection of let-7 into DGCR8$^{-/-}$ mESCs rescued the differentiation defect, allowing the cells to effectively downregulate the pluripotency markers, therefore the let-7 family of miRNAs play a major role in the initiation of differentiation (198).

Large scale transcriptomic studies on DGCR8$^{-/-}$ mESCs, have revealed that there is an upregulation of genes involved in a number of processes including cell cycle

signalling, epigenetic modifications and immunity (8,48), however further protein studies are required in order to accurately validate the individual targets indentified in the study.

Although DGCR8 knockout (6), dicer knockout (7), drosha knockout (199) and argonaute 2 knockout (189) mESCs have been created among other cell lines with impaired miRNA biogenesis, as of yet no cell line that does not express any miRNAs has been created, this is in part due to the abundance of miRNA biogenesis pathways and due to the extensive regulatory reach of miRNAs (2), it is unlikely that these cell lines would be viable, particularly as mice with any one of these genes knocked out are embryonic lethal (200).

5.12. MiRNAs and stem cell metabolism

Notably, there is increasing evidence that miRNAs regulate stem cell metabolism (201,202). In particular the let-7 family of miRNAs, which has been implicated in the regulation of pluripotency(203), has also been implicated in the regulation of the insulin-PI3K-mTOR pathway, a pro-glycolytic pathway (116). Furthermore, Lin28 a protein which is highly expressed in ESCs, and is a negative regulator of the let-7 family has as part of a mix of transcription factors been used to successfully reprogram human fibroblasts (58).

Interestingly, Lin28 overexpression in mESCs results in an abundance of metabolites involved in the threonine catabolism pathway, and overexpression of let-7 miRNAs has the opposite effect, indicating these miRNAs can also play a direct role in regulating the TCA cycle via amino acid metabolism (201).

5.13. Embryonic stem cell cycle (ESCC) regulating miRNAs

It has been estimated that there are 323 distinct miRNA sequences, which make up the 110,000 copies of miRNAs within individual mESCs, of which the majority can be accounted for by six distinct loci, notably four of these loci or their human

homologs have been shown to play roles in cell cycle regulation and oncogenesis (204). This suggests that the distinctive cell cycle of mESCs is a major aspect of their phenotype.

As aforementioned, fourteen key miRNAs that were re-expressed in DGCR8$^{-/-}$ mESCs were able to restore the cell cycle defect (3), of these miR-291a, 291b, 294 and 295 were studied further, as they are highly expressed in mESCs, representing approximately a third of the miRNAs present in mESCs (205,206). miR-291a, 291b, 294 and 295 are part of a cluster (miR-295 cluster) of co-transcribed and co-expressed miRNAs that have been shown to play an important role in controlling the mESC cell cycle, and so have been termed Embryonic Stem Cell Cycle (ESCC) (3). Due to the importance of the miR-295 cluster for the normal mESC phenotype and for embryonic development(3,6), it is highly conserved having homologs in many mammalian organisms including humans (205,206) (Figure 8) and in other organisms such as zebrafish (207), which all have the same seed sequence of 'AAGUGC' (Figure 8). As all of the members of the miR-295 cluster have the same seed sequence (with the exception of miR-291b which has one nucleotide difference), they are functionally redundant and share many of the same targets (3,8,26).

Through transfection studies in DGCR8$^{-/-}$ mESCs, the highly conserved clusters of miRNAs (Figure 8) (205,206) have been shown to target cell cycle inhibitor mRNAs; *Cdkn1a, Lats2* and *Rbl2* (26), which regulate the G_1/S phase transition resulting in restoration of the WT mESC cell cycle (3). Notably, miR-294 was the most efficient at restoring the normal mESC cell cycle (3). Overexpression of the miR-295 cluster in mESCs results in increased resistance to differentiation and sustains their high proliferation rate (208). However, the ESCC miRNAs do not restore the differentiation defects in DGCR8$^{-/-}$ mESCs, indicating that a different subset of miRNAs is important for proper differentiation (3). Interestingly, the ESCC miRNAs do act to prevent mesodermal differentiation by targeting the *Dkk-1* gene, a Wnt pathway inhibitor (209). Notably, the miR-302 cluster (105) has been shown to be directly positively regulated by the pluripotency markers Oct4 and Sox2 (210). This is also the case with the miR-295 cluster which has been

shown to be positively regulated by the three key pluripotency markers; Oct4, Nanog and Sox2 (48). The ESCC miRNAs also inhibit the let-7 family of miRNAs, which are known to promote differentiation (198) and suppress cell cycle progression (211). Therefore the ESCC miRNAs play an essential role in the mESC cell cycle and by extension in the maintenance of the pluripotency phenotype.

As well as promoting the cell cycle and pluripotency, the miR-295 cluster has also been shown to play an important role in protection against apoptosis, studies on Dicer knockout mESCs have shown that there is an upregulation of cell death genes such *Caspase 2* which are targeted by the miR-295 cluster. Notably only when the cells were exposed to DNA damaging chemicals, did the protective properties of the miR-295 cluster become apparent (148). This is supported by murine studies on miR-295 cluster knockout mice which, display partially penetrant embryonic lethality (212). The miR-295 cluster has also been shown to protect against autophagic cell death in human melanoma cells (213).

The role of the ESCC in the maintenance of the pluripotent phenotype has been evidenced by their ability to replace the Yamanaka factors and reprogram mature cells into iPSCs, at a higher efficiency than the more traditional method (16,214). In fact it has been proven that the expression of miRNAs is necessary for reprogramming to occur, as Dicer knockout cells are unable to undergo full reprogramming to iPSCs and instead reprogram into Dicer null mESCs, which have severe growth and differentiation defects, and only upon the re-expression of Dicer and therefore miRNAs, are they able to be reprogrammed (215).

These studies stress the importance of miRNAs and the ESCC miRNAs in particular for the normal mESC phenotype.

5.14. ESC Proteomics

Proteomics is the large-scale study of proteins, with particular emphasis on their structures and functions. In this study we used proteomics to determine which proteins are not only affected by the loss of DGCR8 but also by the presence of the miRNAs known to be highly expressed in mESCs, thereby, providing a detailed and broad overview of potential mechanisms controlling the ES cell phenotype.

Thus far, proteomic studies on ESCs have provided great insights into the proteomes of ESCs both under proliferative and differentiation conditions (216,217). By combining the results of these studies with other studies using transcriptomics (218), the mechanisms underlying the unique phenotype of ESCs can be explored. Therefore the integration of investigative techniques is vital for formulating models of cell behaviour, and may allude to possible regulatory input such as miRNAs. This is most apparent in the few miRNAs studies using -omics techniques, which indicate correlation between decreased protein and RNA expression upon the overexpression of specific miRNAs (216,217).

5.14.1. Sample preparation

The results of any experiment depend greatly on the quality of the starting material, this is particularly important in molecular biology where cellular stress and extrinsic factors such as temperature can have major effects on the quality of the biological material and therefore affect any downstream results (219). All proteomics studies begin with extensive sample preparation to ensure maximum quality of the sample.

The proteins are isolated from cell lysates and solubilised in a buffer. This buffer usually contains numerous denaturing agents including chaotropes and detergents (Table 1).

In order to further disrupt the cell membrane and produce homogenised cell lysates, the cell lysate then undergoes sonication, and centrifugation this allows for the separation of the solubilised total cellular protein and a cell pellet consisting of cellular debris and intracellular molecules such as DNA.

5.14.2. 2D Difference in gel electrophoresis (2D DIGE)

2D DIGE is a quantitative method for comparing whole protein expression, including post-translational modifications. It allows for the routine separation of 1000s of proteins and their isoforms in the same gel (220).

A simplified summary of the 2D DIGE methods is summarised in Figure 10. Firstly, proteins are minimally labelled with fluorescent cyanine dyes (Cy2, Cy3 and Cy5), which bind to surface-exposed lysines allowing for downstream identification of proteins. The Cy3 and Cy5 dyes are incubated with known concentrations of two separate protein samples, and the Cy2 dye is incubated with known concentrations of a pooled sample, in order to act as an internal standard. Excess dye is quenched by the addition of more lysine. Unlike methods that rely amino acid labelling for peptide identification, DIGE is not hindered by the fact that these amino acids may not always be accessible, resulting in proteins being missed (221).

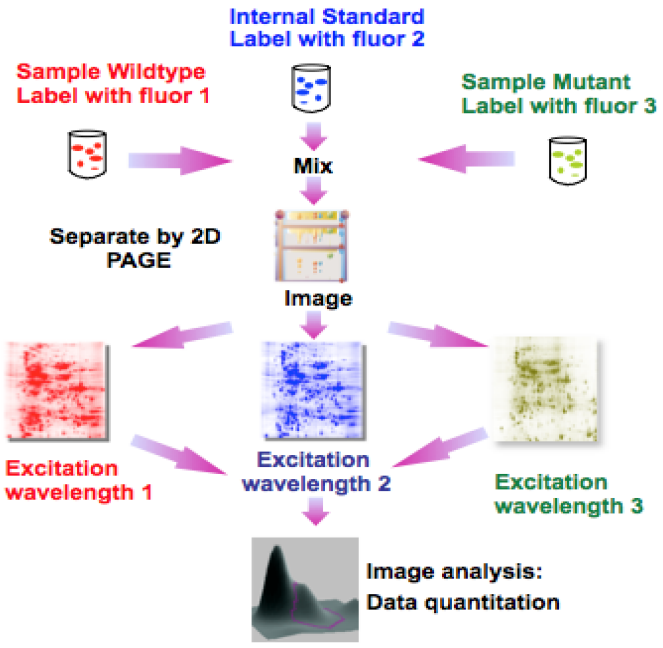


Figure 10 – Summary of 2D Difference in Gel Electrophoresis (2D DIGE) method.

In order to obtain accurate representations of the samples proteomes, samples are loaded multiple times on different gels. Per gel, the Cy3, Cy5 and Cy2 samples are combined and proteins are then separated in the first dimension according to their isoelectric points (pI), the pH at the which the net charge of the protein is zero, this is known as isoelectric focusing (IEF). This is done by loading the proteins onto an immobilised pH strip and applying an electrical field, once proteins are aligned according to their pI, the protein is immobilised (222).

The next step is to separate the proteins in the second dimension by their molecule mass (Mw), in order to do this the strip is equilibrated by saturation in sodium dodecyl sulfate (SDS). The strip is then run on a polyacrylamide gel so that proteins are separated both by their pI and Mw, and this can be used to identify accompanying post-translational modifications (223). Once the proteins have completed running, they are not visible to the naked eye on the gel, therefore the gel is scanned and imported into software (224) in order to view the relative expression levels of the proteins and also to view their positions within the gel for downstream identification. The gel itself can be stained after being fixed in a solution containing acetic or phosphoric acid, which prevents

diffusion of the proteins (220). Conventional staining dyes include Coomasie Blue and Silver nitrate, of which silver nitrate is the more sensitive (225).

Multiple gels can be analysed simultaneously and aligned according to protein/spot position producing a protein 'map'. Spots/proteins of interest can be highlighted on the software over multiple gels, aiding in the selection of spots for mass spectrometry (MS) analysis (Figure 10).

5.14.3. Protein identification

Before MS analysis, the spots on the 2D DIGE gel, are manually picked from the gels and digested using proteolytic enzymes such as trypsin. The enzyme trypsin is a highly specific, high activity, serine protease that cleaves the C-terminal end of lysine (K) and arginine (R) producing tryptic peptides of appropriate lengths for further analysis using MS (226). Next the peptides are analysed using MS based methods.

5.14.4. Ionization

A mass spectrometer works by using magnetic and electric fields to apply forces to ions within a vacuum (227). During ionization, the sample is converted into gas phase ions. Two major methods are primarily used; matrix assisted laser desorption/ionization (MALDI) and electrospray/microspray/nanospray ionization (ESI), these are both similar in that they are referred to soft ionization techniques as they tend to produce mass spectra with little or no fragment-ion content, this in particular makes them optimal for the analysis of high molecular compounds (228). However, MALDI produces ions by pulsed-laser irradiation of a sample whereas ESI produces ions by applying a high voltage to a liquid sample in order to produce an aerosol, this produces various multi charged ions, which allows the analyser to detect peptides over a wide range of masses (229).

5.14.5. Analysis and detection

Once gas phase ions are formed, they are fed into a mass analyser that measures the mass/charge (m/z) ratio of the ions. There are four major types of analysers; time-of-flight, quadrupole, ion trap and Fourier transform ion cyclotron (FT-MS) (227). All of these systems use either an electric and/or magnetic field to affect the path and /or velocity of the ions according to their mass/charge ratios, in order to generate a mass spectrum of the peptides coming from the ion source (227). Often this is coupled to a second MS reaction, whereby targeted ions of a specific m/z are separated from the remaining ions (referred to as parent/precursor ions) and undergo another fragmentation step to generate a new, more targeted mass spectrum, this increases the specificity of detection of known peptides (227). There are a number of methods that are used for the fragmentation step including electron transfer dissociation (ETD), collision-induced dissociation (CID) and electron capture dissociation (ECD) (230). The final step in mass spectrometry is detection, this occurs via a detector which records either the charge or current produced when an ion passes or hits it surface producing a mass spectrum (230).

The proteins are then identified by aligning the mass spectrum with comprehensive protein databases such as Uniprot (231) or NCBI. Once the proteins have been successfully identified, the Mw and pI of the proteins provided by the databases can be compared to the Mw and pI of the proteins/spots on the DIGE gel, to ensure that the proteins are not correctly matched up but to also highlight the presence of post translational modifications.

The methodology of mass spectrometry is summarised in Figure 11.

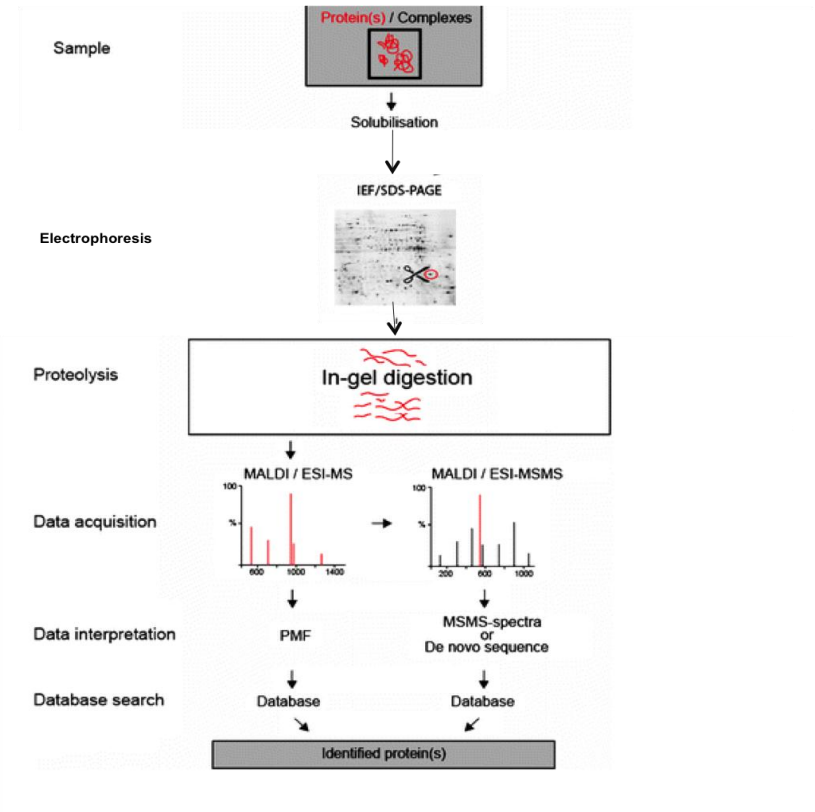


Figure 11 – Summary of Mass Spectrometry on a 2D-DIGE gel Adapted from: (232)

5.15. Proton nuclear magnetic resonance (H^1 NMR) spectroscopy

H^1 NMR is technique that harnesses the magnetic properties of nuclei to identify molecules, in the case of H^1 NMR, it focuses on particularly on molecules that contain hydrogen. It works because specific nuclei in a magnetic field have the ability to absorb and reabsorb at specific electromagnetic radiations allowing researchers to identify them. Due to the nature of the technique using hydrogen atoms as a way of identifying compounds, it is imperative that samples are pH'ed and adjusted to pH 7 so as to not interfere with any readings. In this study we used H^1 NMR, to not only study the secretomes of the WT and DGCR8$^{-/-}$ mESCs, but also to study the intracellular metabolite composition of the cells.

In order to extract soluble metabolites from the mESCs, in a way that minimised any alterations to metabolite concentration in the cells due to factors such as cellular stress, or metabolite degradation, immediately prior to harvest the cells

were rapidly cooled down to restrict enzymatic degradation and then they were harvested in perchloric acid (233).

Our group has previously demonstrated that the combination of proteomics and metabolomics offers complementary information (234).

6. Hypotheses

miRNAs are known to display their effects at the protein level, yet large scale studies focusing on miRNA function in mESCs predominantly use transcriptomic techniques to study their targets and functions (8). As the major readout of miRNA function is at the protein level, this provides an incomplete 'picture' of the targets of these miRNAs, and their role in the mESC phenotype.

1. Pathways governing mESC differentiation are controlled by miRNAs. A proteomic approach comparing differential protein expression in DGCR8 knockout mESCs and WT mESCs grown in the presence of LIF will reveal miRNA-dependent proteins that influence pluripotency.

2. DGCR8 knockout mESCs provide a relatively 'blank canvas' for the study of individual miRNA function. A proteomic approach comparing differential protein expression in DGCR8 knockout mESCs transfected with individual ESCC miRNAs grown in the presence of LIF, will reveal direct and indirect targets that are shared between the different miRNAs and affected by the individual miRNAs and potentially reveal their associations with pluripotency.

7. Aim

Before stem cell-based therapies can be used for treatment, it is essential that we gain knowledge of the molecular mechanisms controlling pluripotency and differentiation towards the cardiovascular lineage. Because of their ease of expansion, allowing a large number of cells to be obtained, and the existence of the DGCR8$^{-/-}$ mESC cell line, mESCs will be used in the present study.

7.1. Analyses comparing WT and DGCR8 knockout mESCs

i) A comprehensive proteomics analyses of DGCR8 null mESCs and wild type mESCs.
ii) Identify and validate miRNA-dependent proteins in mESCs.
iii) Study the effects that the loss of DGCR8 has on the metabolome of mESCs

7.2. Analyses comparing DGCR8 null mESCs transfected with the individual ESCCs

iv) A comprehensive proteomics analyses of DGCR8 knockout mESCs transfected with the ESCC miRNAs
v) Identify direct/indirect and validate targets of the ESCC miRNAs in DGCR8 knockout mESCs
vi) Study the effects that the individual ESCCs have on the DGCR8 knockout mESCs on both the metabolome and proteome of the cells.
vii) Compare the proteins identified in both of the proteomics analyses (i and iv), to identify common targets.

8. Methods

8.1. Cell culture

All cells were grown in an incubator, at 37° C in a humidified atmosphere of 5% CO_2.

8.1.1. Derivation of MEF feeder layer

The clonal mouse embryonic fibroblast (MEF) cell line (SCRC-1008, ATCC) derived from C57BL/6 mouse embryos, was removed from liquid nitrogen and thawed rapidly by swirling in a 37°C water bath. The cell suspension was then transferred into a falcon tube and resuspended in complete MEF medium before being centrifuged for 5 minutes at 1000 rpm. Next, the supernatant was discarded and the cell pellet was resuspended in 5ml of medium and plated in a sterile T25 cell culture flask. The following day the medium on the cells was changed to fresh medium. Complete MEF medium consisted of DMEM (ATCC, Cat No. 30-2002) supplemented with 15% FBS (ATCC, Cat No. SCRR-30-2020) 100 U/ml Penicillin and 100mg/ml Streptomycin (Life Technologies, Cat No. 15140-122) and 5ml of 200mM L-glutamine (Life technologies, Cat No. 25030-024). The complete medium was allowed to reach room temperature before use.

Once the cells were ≥90% confluent, they were subcultured: the medium was removed and, the cells were washed twice with DMEM (Life technologies, Cat No. 10938025) supplemented with 5ml of100 U/ml Penicillin and100mg/ml Streptomycin (Life Technologies, Cat No. 15140-122). Next, 3ml of trypsin-EDTA solution (Life technologies, Cat No. 25300-054) was added to the flask and incubated for 90 seconds at 37°C or until cells began to detach from the flask. The trypsin-EDTA solution was then neutralized by the addition of 4 ml of complete MEF medium and the cells were transferred to a sterile falcon tube and centrifuged at 1000rpm for 5 minutes at room temperature.

Subsequently, the supernatant was removed from the cells, and the cells were resuspended in 5 ml of complete MEF medium, and plated at a ratio ranging from

1:3 to 1:5 in T25s. Medium was changed every 1-2 days. Once the cells had reached 90% confluency, they were treated with 12µg/ml of mitomycin C (Sigma, Cat No.M4287) (2mg of mitomycin C was dissolved in 5ml of ddH_2O to create a 0.4mg/ml stock solution of which, 30µl of mitomycin C per 1 ml of medium was added to the cells) in complete MEF medium for 3 hours at 37°C. Next, the cells were washed several times with DMEM (Life technologies, Cat No. 10938025) supplemented with 5ml of 100 U/ml Penicillin and 100mg/ml Streptomycin (Life Technologies, Cat No. 15140-122) before being incubated overnight at 37°C in complete MEF medium. Prior to use as a feeder layer, the cells were washed a further three times with DMEM (Life technologies, Cat No. 10938025) supplemented with 5ml of 100 U/ml Penicillin and 100mg/ml Streptomycin (Life Technologies, Cat No. 15140-122).

Feeder layers were not maintained for longer than one week and only used up to passage 7 to retain optimal function.

8.1.2. Mouse embryonic stem cells

WT (v6.5) and DGCR8$^{-/-}$ mouse embryonic stem cells (Novus Biologicals, Cat nos. NBP1-41162 & NBA1-19349, respectively) were removed from storage, thawed rapidly by swirling in a 37°C water bath, resuspended in complete mouse embryonic stem cell (mESC) medium and centrifuged at 1000rpm for 5 minutes. Next the supernatant was removed and the cells were resuspended in 5ml of complete mESC medium before being plated on a pre-treated MEF cell feeder layer (DGCR8$^{-/-}$ mESCs). v6.5 cells were plated directly onto 0.1% gelatin (Sigma, Cat No. G-1393) (8 ml of 2% gelatin was combined with 160ml of DMEM (Life technologies, Cat No. 10938025)) coated T25 flasks and passaged onto gelatin-coated flasks. The mESC complete medium consisted of DMEM (ATCC, Cat No. 30-2002) supplemented with 15% FBS (ATCC, Cat No. SCRR-30-2020), 5ml of 100 U/ml Penicillin and100mg/ml Streptomycin (Life Technologies, Cat No. 15140-122). 100x non essential amino acids (Life technologies, Cat No. 11140-050), 5ml of 200mM L-glutamine (Life technologies, Cat No. 25030-024), 0.1mM

2-mercaptoethanol (BDH, Cat No. 441433A) and 10ng/ml recombinant leukaemia inhibitory factor (LIF; Chemicon, Cat No. LIF1010).

The next day the medium was changed and once the cells were 70-80% confluent, they were passaged in a similar manner to the protocol used for the MEF cells. However, after the DGCR8$^{-/-}$ cells were centrifuged, they were subcultured into a ratio ranging from 1:2 to 1:5 and plated in 0.1% gelatin coated T25 culture flasks for 30 minutes at 37°C in order to allow for a separation of the mESCs and the feeder layer. Next, the supernatant containing unattached cells (predominantly DGCR8$^{-/-}$ cells) was removed, and 4 x 10^5 DGCR8$^{-/-}$ cells were plated on newly treated MEF cells in T25s. The WT cells were plated directly onto 0.1% gelatin coated plates. Cells were subcultured every 2-3 days.

Before experiments were conducted on DGCR8$^{-/-}$ mESCs, cells were weaned off the mitomycin C treated MEF layer and grown in gelatin-coated flasks. Similar to the subculturing procedure, cells were washed, trypsinized, centrifuged, and plated onto 0.1 % gelatin coated plates for 30minutes. Next, the number of cells in the supernatant were counted and the appropriate number of cells were transferred to a new gelatin coated flask, after 2-3 days, cells were passaged directly onto 0.1% gelatin coated flasks. Cells were passaged at least three times onto gelatin-coated flasks to eliminate MEF contamination.

8.1.3. Alkaline phosphatase staining

DGCR8$^{-/-}$ and WT cells were plated at a density of 25,000 and 50,000, respectively, in a 0.1% gelatin-coated 6 well plate. Medium was changed every two days and staining was conducted using an alkaline phosphatase detection kit (Millipore, Cat No. SCR004) on the fifth day of culture, as per manufacturers' instructions. Briefly, media was removed from cells and they were fixed in 4% paraformaldehyde in PBS for 90 seconds. Cells were washed with PBS (Sigma Aldrich, Cat No. D8537) with 0.1% Tween (Sigma Aldrich, Cat No.P1379) and covered in staining solution in the dark, at room temperature for 15 minutes.

Cells were washed again with PBS-Tween, covered in PBS and stained colonies were observed using a light microscope.

8.1.4. Cell transfection

DGCR8$^{-/-}$ cells were transfected with the Embryonic Stem Cell Cycle (ESCC) miRNAs in mimic form, four from the miRNA 290 cluster; miR-291a-3p (Life Technologies, Cat No. mc12953) miR-291b-3p (Life Technologies, Cat No. mc13011), miR-294-3p (Life Technologies, Cat No. mc10865), miR-295-3p (Life Technologies, Cat No. mc10386) and miR-302d-3p (Life Technologies, Cat No. mc10927), as well as a negative control miRNA (Qiagen, Cat No. 1027280). Twenty four hours before the cells were transfected, the DGCR8$^{-/-}$ cells were plated in complete medium at the densities shown in Table 2.

Prior to transfection, two mastermixes labelled A and B (Table 3) were made up as shown in transfection media consisting of DMEM (ATCC, Cat No. 30-2002) and 10ng/ml recombinant LIF (Chemicon, Cat No. LIF1010).

Mastermix A and the corresponding Mastermix B were mixed in one tube by gentle pipetting, and incubated at room temperature for 10 minutes before being added to the correct flask.

Meanwhile the medium was removed from the cells to be transfected, and the cells were washed three times with DMEM (Life technologies, Cat No. 10938025) supplemented with 5ml of 100 U/ml Penicillin and100mg/ml Streptomycin (Life Technologies, Cat No. 15140-122). Next 6 ml of transfection medium was added to the cells in the T75 flasks and, 3ml of transfection medium was added to the cells in the T25 flasks. Once the mastermixes had been added to the corresponding flasks, the flasks were placed in the incubator for five hours. Subsequently, 6ml and 3ml of neutralizing medium consisting of DMEM (ATCC, Cat No. 30-2002) supplemented with 30% FBS (ATCC, Cat No. SCRR-30-2020) and 10ng/ml LIF (Chemicon, Cat No. LIF1010), was added to the T75s and T25s, respectively. The next morning, the medium on the cells was changed to

complete medium and the cells were incubated for a further 48 hours before harvesting.

Three independent transfection experiments were performed per condition.

8.2. RNA

8.2.1. RNA extraction

Cells were washed twice with PBS and the T25 flasks were turned on their sides to drain cells of all PBS. RNA extraction was conducted using the miRNeasy Mini Kit (Qiagen, Cat No. 217004) following the manufacturers' protocol. Briefly, cells were scraped in 700µl of Qiazol Lysis Reagent into 1.5 ml eppendorf tubes, incubated at room temperature for 5 minutes, and then vortexed with 140µl of chloroform. The lysate was incubated for a further 5 minutes at room temperature before being centrifuged for 15 minutes at 12,000 rpm, at 4°C. 280 µl of the upper aqueous phase was transferred to a new tube, mixed with ~ 420µl of 100% ethanol by thorough pipetting and transferred into an RNeasy mini column. The remaining centrifugation steps were carried out at room temperature. The column was centrifuged for 1 minute at 13,000 rpm and the flowthrough was discarded, next 700µl of RWT buffer was added to the column, it was centrifuged for 1 minute at 13,000rpm and the flow through was discarded. Next 500µl of RPE buffer was added to the column before a further centrifugation step for 1 minute at 13,000rpm, the flow through was discarded and a further 500µl of RPE buffer was added to the column before it was centrifuged at 13,000rpm for 2 minutes. At this point, the mini column was transferred into a new collection tube and centrifuged at high speed for 1 minute, before being transferred to a new tube. The RNA was eluted in 25-30µl of RNase free water, by centrifugation at high speed for 1 minute. RNA concentration was measured using the nanodrop at 260nm.

8.2.2. Reverse transcription PCR prior to conventional PCR

For each PCR reaction, 2.5µg of RNA was used, the total volume was adjusted to 10µl with nuclease free water. The RNA was then reverse transcribed into cDNA using reagents from the Improm-II Reverse Transcription kit (Promega, Cat No. A3800); the 5x Reaction buffer, $MgCl_2$ (25mM), Random Primers, dNTPs, RNasin and the reverse transcriptase enzyme, as well as dNTPs (25mM) (Life Technologies, Cat No. 10297-018).

The reverse transcription reaction was set up by mixing the 10µl of RNA with 0.2µl of Random Primers, 5µl of 5x Reaction buffer, 3µl of $MgCl_2$ (25mM) and 1.25µl of 25mM dNTPs. The samples were then placed in a thermal cycler, where the temperature was raised to 70°C for 5 minutes, and then dropped to 4°C for 5 minutes whilst 0.625µl of RNasin, 3.925µl of RNase free water and 1µl of reverse transcriptase was added to each sample, making the final reaction volume 25µl. The RT-PCR reaction program was as follows, samples were initially kept at 25°C for 5 minutes to allow the primers to anneal, then reverse transcribed for 90 minutes at 42°C, and finally the reverse transcriptase enzyme was inactivated for 15 minutes at 72°C. Samples were then diluted to 10ng/µl using RNase free water.

8.2.3. Polymerase chain reaction (PCR)

Primers were designed (Table 4) and cDNA was amplified using a PCR kit (Life Technologies, Cat No. 18038042). PCR reagents were mixed on ice following the manufacturers' instructions; to 5µl of 10ng/µl cDNA, a mixture of 1.5µl of 10x PCR buffer, 0.6µl of 50mM $MgCl_2$, 0.1µl of 25mM dNTPs, 2µl of the appropriate forward and reverse primers (Table 4) at a concentration of 10mM each and 5.7µl of RNase free water was added. Tubes were briefly centrifuged and placed in a thermocycler for the following program; preheat lid at 105°C, initial denaturation at 94°C for 4 minutes, then 30 cycles of denaturing of the cDNA at 94°C for 1 minute, annealing of the primers at 58°C for 1 minute and elongation

at 72°C for 1 minute. This was followed by a final extension step at 72°C for 5 minutes.

Samples were run on a 2% agarose gel with Safeview (NBS Biologicals, Cat No. NBS-SV1).

8.2.4. Reverse transcription prior to quantitative PCR for genes

200ng of RNA was topped up to 4.5µl with RNase free water. The RNA was then reverse transcribed using the High Capacity RNA to cDNA kit (Life Technololgies, Cat No. 4387406), containing 20x Enzyme mix and 2x RT buffer mix.

The reverse transcription reaction was set up by mixing 4.5ul of the RNA with 0.5µl of the enzyme mix and 5µl of the RT buffer mix. The samples were then placed in the thermocycler for the following program, 16°C for 5 minutes to allow the primers to anneal, 37°C for 60 minutes to allow the reverse transcription reaction to occur and 95°C for 5 minutes to inactivate the reverse transcriptase enzyme.

8.2.5. Reverse transcription prior to quantitative PCR for miRNAs

For each reverse transcription reaction, 200ng of RNA was used in a final volume of 3µl (topped up with RNAse free water). The RNA was then reverse transcribed using reagents from the Taqman MicroRNA Reverse Transcription Kit (Life Technologies, Cat No. 4366596), containing 100mM of dNTPs with dTTP, 10x RT buffer, $MgCl_2$ (25mM) and RNase Inhibitor (20Uµl). In addition the Megaplex RT Primers, Rodent Pool A (Life Technologies, Cat No.4399970) and Multiscribe Reverse Transcriptase (Life Technologies, Cat No. 4311235) were added.

The reverse transcription reaction was set up by mixing the 3µl of RNA with 1µl of the Megaplex RT primers, 0.3µl of dNTPs, 2µl of the Reverse Transcriptase, 1µl of the 10x RT buffer, 1.2µl of $MgCl_2$, 0.2µl of RNase Inhibitor and 1.3µl of Nuclease free water. The samples were then placed in a thermocycler, for a

program where for 40 cycles, the temperature was held at 16°C for 2 minutes, raised to 42°C for 1 minute and then raised to 50°C for 1 second. Finally the samples were held at 85°C for 5 minutes. Samples were then diluted to 2ng/ul with RNase free water.

8.2.6. Quantitative PCR (qPCR) using taqman assays

Samples were loaded onto a 384 well plate. Per reaction, 1µl of 2ng/µl cDNA was mixed with 2.5µl of Taqman PCR Mastermix (Applied Biosystems, Cat No. 4440049), 0.25µl of the appropriate Taqman microRNA expression assay or gene expression assay (Table 5) and 1.25µl of RNAse free water. Each reaction was carried out in duplicate.

8.2.7. Quantitative PCR (qPCR) using SYBR green

Samples were loaded onto a 384 well plate. Per reaction, 1µl of 2ng/µl cDNA was mixed with 5µl of SYBR® Select Master Mix (Life Technologies, Cat No. 4472908), 1µl of the appropriate primer pair (Table 4) at a concentration of 10µM each and 3µl of RNAse free water.

8.3. Proteomic analyses

8.3.1. Cell protein extraction

Cells were washed twice with chilled DIGE wash buffer (100mM, pH 8 Tris with 1 M Magnesium acetate) and scraped in 1 ml of DIGE wash buffer, before being centrifuged at 4°C, at high speed for 2 minutes. The supernatant was then removed and the cell pellet was resuspended in 150µl of DIGE lysis buffer (3ml of 1M Tris, 48g of Urea, 4g of CHAPS and 2 Protease inhibitor tablets (Roche, Cat no. 06538304001), before being sonicated for 15 seconds twice at 4°C. Cell lysates were incubated on ice for 30 minutes, with a brief vortex at 5 minute intervals, before being centrifuged at maximum speed, at 4°C for 10 minutes. The supernatant (containing all cellular proteins) was retained for further experiments.

8.3.2. Protein concentration measurement

Protein concentration was measured using the Biorad Protein Assay Reagent (Biorad, Cat No. 500-0006) following the manufacturer's protocol. Briefly, the Bradford dye was diluted in dH_2O in the ratio 1:4. A BSA (bovine serum albumin) standard stock solution was made up at the concentration 10mg/ml with dH_2O and serial dilutions were prepared. For measurements, 2µl of cell lysate or 1µl of the standard mixed with 2µl of the DIGE lysis buffer was mixed with 998µl of the diluted dye, and incubated for 5 minutes before the optical density was measured at 595nm. Using the optical density measurements the protein concentrations were determined using the BSA standard curve.

8.3.3. 2D DIGE clean up

2D DIGE clean up of samples was carried out using the ReadyPrep 2-D Cleanup Kit (Bio-rad, Cat No.163-2130), following the manufacturers' protocol. All steps were carried out on ice. Briefly, 500µg or 100µl of each protein was transferred into an individual standard eppendorf tube, and 300µl of precipitating agent was added. The solution was then briefly vortexed and incubated on ice for 15 minutes. Next, 300µl of precipitating agent 2 was added to the mixture, and mixed by vortexing. The samples were then centrifuged at maximum speed for 5 minutes and the supernatant was discarded, before 40µl of wash reagent 1 was added to the pellet and it was again centrifuged at maximum speed for 5 minutes. The wash reagent was then removed, and 25µl of ultrapure water was added to the pellet, before 1 ml of pre-chilled wash reagent 2 and 5µl of wash additive 2 were added to the pellet, and the pellet was vortexed for 1 minute. The samples were then incubated at -20°C for 30 minutes, and vortexed every 10 minutes for 30 seconds. After the incubation, the samples were centrifuged at maximum speed for 5 minutes, the supernatant was discarded and the pellets were air-dried for 5 minutes. Pellets were resuspended in 100µl of DIGE lysis buffer.

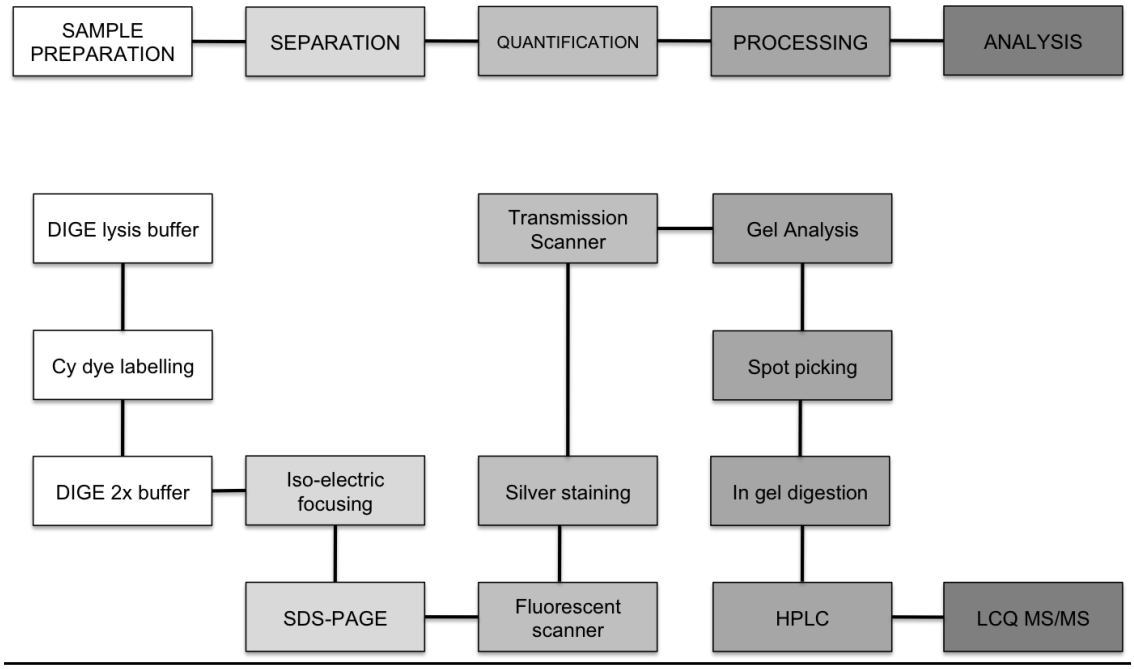


Figure 12 - Summary of the proteomics workflow

The fluorescence dye reaction was carried out at a dye/protein ratio of 200pmol/50µg. 50µg of each sample was labelled with either Cy3 or Cy5 dye dissolved in DMF (Sigma Aldrich, Cat No. 227056) for direct comparison of the two samples, and 25µg of each sample was labelled with Cy2 as an internal standard control. Labelled samples were incubated at 4°C in the dark for 30 minutes, then 10mM of lysine (Sigma Aldrich, Cat No.L8662) was added to each sample for 15 minutes. Per gel, samples were labelled with the CyDye DIGE Fluor Minimal labelling kit (GE Healthcare, Cat No. GE25-8010-65), Cy3, Cy5 and Cy2 were mixed and 2x DIGE buffer (8M urea, 2% v/v Pharmalyte, 2% w/v DTT and 4% CHAPS) was added, samples were topped up to 450µl with rehydration buffer (8M urea, 0.5% w/v CHAPS, 0.2% w/v DTT, 0.2% w/v Pharmalyte and trace amounts of bromophenol blue). Protein samples (two samples and one control per strip) were loaded onto 18cm non-linear immobilized pH gradient strips, pH 3-10 (GE healthcare, Cat No. 17-6003-76) on a reswelling tray, covered with Immobiline DryStrip cover fluid (GE Healthcare, Cat No. 17-1335-01) and left in the dark, overnight at room temperature. Strips were focused in the Multiphor™ II isoelectric focusing system (GE healthcare) at 20°C with the following protocol: 150V for 2 hrs, 300V for 2 hrs, 600V for 2 hrs, 1500V for

8hours, 8000V for 30mins, then 8000V for 2 hrs. Once the isoelectric focusing had completed, the strips were equilibrated in equilibration buffer (6M urea, 2% w/v SDS, 30% glycerol, 50mM Tris/HCl pH 8.8 and trace amounts of bromophenol blue) with 1% w/v DTT for 15 minutes on a shaker, followed by a further 15 minute incubation with equilibration buffer with 4.8% iodoacetaemide on a shaker. The strips were briefly washed with water and loaded onto 12%T (total acrylamide concentration), 2.6% C (degree of crosslinking) polyacrylamide gels without a stacking gel. The gels were then sealed with an agarose sealing solution (100ml of SDS electrophoresis buffer, 1g of agarose and trace amounts of bromophenol blue), before being run in the Ettan™ DALT six vertical electrophoresis system (GE Healthcare). The gels were run at 10°C with this protocol; for 6 gels, 600V, 400mA, 12W for 15 mins, 15W for 30 mins and 100W for approximately 4.5 hrs, using 2x running buffer (Life Technologies, Cat No.LC2675) in the upper chamber and 1x running buffer in the lower chamber, until the bromophenol blue dye had migrated off the lower end of the gels. Fluorescence images were acquired using the Typhoon variable mode imager 9400 (GE Healthcare).

8.3.4. Silver staining

Protein profiles of 2-DE gels were visualized by silver staining using the PlusOne™ Silver Staining Kit (GE Healthcare, Cat No.17-1150-01) and protocol with slight modifications, for compatibility with mass spectrometry. All steps were performed with gentle shaking and all solutions were made with double distilled water. Gels were fixed in fixing solution (40% v/v methanol, 10% acetic acid) overnight. Next gels were sensitized for 30 minutes in sensitizing solution (30% v/v methanol, 0.2% w/v sodium thiosulphate and 0.5M sodium acetate), followed by 3 x 5 mins washing steps in water and silver staining for 20 minutes in 0.25% w/v $AgNO_3$. After two further 5 min washes, the developing solution (2.5% w/v sodium carbonate, 0.0148% w/v formaldehyde) was added and the gels were gently shaken until the silver spots became visible but before the background staining became too bright. At this point, the developing solution was replaced by the stopping solution (1.46% w/v EDTA-Na_2•$2H_2O$) for 10 minutes. Next, the gels were washed 3 x 5 mins with water and placed in clear

plastic bags to be scanned in transmission scan mode using a calibrated scanner (GS-800, Bio-Rad).

8.3.5. Analysis of 2-DE gels

Gel images were analysed using the DeCyder software (GE Healthcare), and protein spots showing a statistically significant difference i.e a fold change of 1.2-1.5, p value<0.05, were selected for manual picking.

8.3.6. In gel tryptic digestion

Using the information from the DeCyder software, spots were picked from at least 4 of the total number of gels. In gel digestion was performed using an Investigator ProGest robotic digestion system (Genomic Solutions). Briefly, spots were destained with destaining solution (15mM potassium ferricyanide, 50 mM sodium thiosulfate) for 15 mins, washed with 100µl of double distilled water for 2 x 10 mins, then with 100µl of 25mM ammonium bicarbonate (Sigma Aldrich, Cat No. A6141). All solutions were purged using Nitrogen gas, gels were dehydrated with 50µl of acetonitrile (Sigma Aldrich, Cat No. 27071-7), and neutralised with ammonium bicarbonate, before being shrunk with 100µl of acetonitrile. The gels were reduced in 30µl of 10mM DTT in 50mM ammonium bicarbonate at 60°C before being cooled for 20 minutes and then alkylated in 30ul of 50mM iodoacetamide in 50mM ammonium bicarbonate for 15 mins. After a 10 minute wash with 40µl of 50mM of ammonium bicarbonate, gel pieces were shrunk with 50µl of acetonitrile for 2 x 15 mins and 15µl of trypsin solution (10µg of trypsin in 600µl of 2mM HCl in 10% acetonitrile, and 900µl of 25mM ammonium bicarbonate) was added to each well, the digestion was carried out at 37°C for 7.5 hours. 1.5 hours later, 10µl of double distilled water was added to each well.

Subsequently, 10µl of 25mM ammonium bicarbonate, 20µl of acetonitrile and 20µl of formic acid was added sequentially to the wells, for 10 minutes each.

Next 20µl of acetonitrile was added to each well for 15 minutes and 30µl more was added for a further 15 minutes, before all the solutions were purged into the collection plate and the digestion was complete. Finally, the gel pieces were lyophilized and resuspended in 18µl of 0.1% formic acid in preparation for mass spectrometry.

8.3.7. Mass spectrometry

Following enzymatic digestion, tryptic peptides were separated on a nanoflow high-performance liquid chromatography (HPLC) (Acclaim®, Thermo Scientific) with subsequent analysis by tandem mass spectrometry (MS/MS) (Q Exactive Plus Orbitrap mass spectrometer, Thermo Scientific). For peptide separation on HPLC reversed phase column was used (PepMap C18, 3µM, 100 Å 50cm x 75µm) and eluted with a 240 min gradient and a flow of 300 nl/min (2-10% B from 0-10 min, 10-30% B from 10-200 min, 30-40% B from 200-210 min, 99% B from 210-220 min and eventually 2% B from 220-240 min; the mobile phase solvent compositions were as follows: A = 0.1% formic acid (FA) in HPLC H_2O; B = 80% acetonitrile, 0.1% FA in HPLC H_2O). The sequentially eluted peptides were then directly analyzed by an orbitrap mass analyzer using full ion scan mode over the mass-to-charge (m/z) range of 350 – 1600, resolution 70000 (at m/z 200). MS/MS was performed using higher energy collisional dissociation (HCD) on the 15 most abundant ions (top 15 ions) in each full MS scan with dynamic exclusion (30sec).

Data from the mass spectrometer was uploaded to the Mascot server and the Scaffold3 program with the parameters allowing for one missed cleavage per peptide, carbamidomethylation of cysteines, and partial oxidation of methionine. This allowed for accurate identification of the differentially expressed proteins based on identified peptides and level of expression.

8.3.8. Buffer exchange

Buffer exchange was carried out on DIGE samples before they were used for immunoblotting, using the Amicon Ultra-0.5 Centrifugal Filter Devices (Millipore, Cat No. UFC500396), following the manufacturers' protocol. Briefly, the filter device was slotted into the provided centrifuge tube and 20-50µl of the protein sample was loaded into the filter device. This was diluted 10x with double distilled water, and the filter device was capped and centrifuged at 14,000 g for 25 minutes, or until only 20-50µl of liquid remained in the filter device. Next, the filter device containing 20-50ul of the protein sample diluted in water was turned upside down into a new eppendorf tube and centrifuged at 1000 g for two minutes.

Protein concentration was measured using the Biorad Protein Assay Reagent (Biorad, Cat No. 500-0006) following the manufacturer's protocol. Using the optical density measurements the protein concentrations were determined using a BSA standard curve.

8.3.9. Immunoblotting

20-30µg of cell protein lysate was mixed with protein loading buffer (0.1mM Tris, 0.1% of 1M SDS, 10% glycerol, 0.25% beta-mercaptoethanol and trace amounts of bromophenol blue) and denatured at 96°C for 10 minutes before being briefly centrifuged and loaded onto a 4-12% Tris-Glycine pre-cast gel (Life Technologies, Cat No. EC60385BOX) or a 12% Tris-Glycine pre-cast gel (Life Technologies, Cat No. EC60055BOX) in NuPage Mops SDS Running Buffer 20x (Life Technologies, Cat No. NP0001) diluted with distilled water. Samples were run at 160v for ≈70 minutes, before being transferred to nitrocellulose membranes (GE healthcare, Cat No. RPN2020D) in ice-cold transfer buffer (25mM Tris Base (3g) and 200mM Glycine (14g) dissolved in 800ml deionized water and 200ml Methanol). The membranes were then blocked in 5% milk in PBS (PBS, Lonza, Cat No. 17-517Q) with 0.1 %Tween for one hour on a shaker at room temperature, briefly washed in PBS-Tween and then incubated with the

appropriate primary antibody (Table 6) diluted in 5% BSA (Sigma Aldrich, Cat No. A2153) with 0.01% sodium azide (Sigma Aldrich, Cat No. S2002) overnight at 4°C on a shaker.

The next day, the membrane was washed 3 times for 15 minutes with PBS-Tween on a shaker and then incubated with the appropriate secondary antibody (Table 6) diluted in 5% milk dissolved in PBS-Tween at a dilution of 1/2000 for 90 minutes at room temperature. The membrane was then washed a further four times with PBS-Tween for 15 minutes each time before ECL Western Blotting Detection Reagent (GE Healthcare, Cat No. RPN2209) was applied for 1 minute and proteins were detected using X-ray films (FUJIFILM, Cat No.AUT-300-040D).

8.4. Cell death (apoptosis) assay

The cell death detection assay was carried out using the Cell Death Detection ELISA^{PLUS} kit (Roche Applied Science, Cat No. 11774425001) per the manufacturers' instructions, the optimal number of cells required for the assay was determined by a titration experiment. 5,000 DGCR8$^{-/-}$ and 2,500 WT cells were plated in a gelatin coated 96 well plate in 100µl of complete mESC medium in triplicate. 48 hours after plating, the medium was removed, 200µl of the provided lysis buffer was added to each well and the plate was incubated at room temperature for 30 minutes. The plate was then centrifuged at 200 x g for 10 minutes and 20µl of each cell lysate, the provided positive control and background control were pipetted into individual wells of the provided streptavidin coated microplate. 80µl of the Immunoreagent was added to each well, and the plate was placed on a shaker at 300rpm for 2 hours at room temperature, before each well was washed three times with Incubation buffer. Next 100µl of the ABTS solution was added to each well, and the plate was placed onto a shaker at 250rpm for approximately 15 minutes or until the green colour had developed enough for proper analysis. Absorbance was measured at 405nm, reference wavelength; 490nm. Absorbance values were normalized to cell number, and were directly proportional to the degree of cell death.

8.5. Metabolomics

8.5.1. Perchloric acid (PA) extraction of cells

PA extraction was undertaken on cells grown in T25s. The medium was removed from the cells, and immediately placed on ice, before being centrifuged for ten minutes at 5000g to pellet any cellular debris. The supernatant was then transferred to a new tube and placed at -80°C until it could be analysed. The cells were washed twice with 10ml of saline, which was removed completely after each wash. Subsequently 4ml of ice cold PA (Sigma Aldrich, Cat No. 244252) was added to each flask, and tilted back and forth to ensure that all the cells were covered. Next the cells were scraped into the PA, and the lysate was transferred into a clean tube on ice. The flask was further washed with 1 ml of PA to ensure that all cells had been collected and the final wash was added to the remainder of the lysate. The cell lysate and PA residue were centrifuged for 15 minutes at 10000rpm for 15 minutes. The supernatant was removed and placed in a fresh tube, to be pH'ed and neutralised to pH 7 with the addition of 1% potassium hydroxide, 10% potassium Hydroxide and 1% PA. The cell pellet was immediately stored at -80°C. Once the supernatant had been pH'ed it was frozen at -80°C, freeze dried and then stored at -80°C again before it underwent metabolic analysis.

8.5.2. Proton nuclear magnetic resonance (H^1 NMR)

H^1 NMR was carried out by Dr. Yuen Li Chung of the Institute of Cancer Research on medium taken after 48 hours of cell growth as previously described.

8.6. Seahorse bioscience

The Seahorse bioscience XF^E24 analyzer was used to measure the Oxygen Consumption Rate (OCR) and Extracellular Acidification Rate (ECAR) of both WT cells and $DGCR8^{-/-}$ mESCs, according to the manufacturers' instructions. Briefly,

the provided 24 well microplates were coated with 0.1% gelatin, and 10,000 WT and 15,000 DGCR8$^{-/-}$ mESCs were plated in triplicate into each well, to ensure approximately 90% confluence at the time of analysis. Four of the twenty wells did not contain any cells acting as a background control. All cells were grown in complete mESC medium for 24 hours, before being washed 3 times with DMEM (Life technologies, Cat No. 10938025) supplemented with 5ml of100 U/ml Penicillin and100mg/ml Streptomycin (Life Technologies, Cat No. 15140-122), and placed in 450µl of complete mESC media either in the presence or absence of LIF for a further 48 hours before mitochondrial function was studied using the Seahorse Bioscience Analyser.

The afternoon prior to analysis, a separate plate containing the drugs ports and probes was rehydrated using the supplied XF Calibrant solution (Seahorse Bioscience, Cat No. 100840-000), this was carried out by pipetting 1ml of the calibrant solution into each well, putting the lid containing the probe and drug ports on top, sealing the plate with parafilm and placing it in a 37°C incubator without CO_2. The next day, both the unsupplemented XF media (Seahorse Bioscience, Cat No. 102365-100) and the XF media supplemented with 25mM glucose (Sigma Aldrich, Cat No. G7021) and 2mM pyruvate (Sigma Aldrich, Cat No. P5280) were warmed to 37°C and adjusted to pH 7. The cells were carefully washed with the unsupplemented media twice, before 450µl of the supplemented media was added to each well. The cells were then placed in an incubator without CO_2 for approximately 45 minutes. Meanwhile, individual stock solutions of 10µM of oligomycin A (Sigma Aldrich, Cat No.75351), 5µM of FCCP (Sigma Aldrich, Cat No. C2920), 10µM of antimycin A (Sigma Aldrich, Cat No. A8674) and 10µM of Rotenone (Sigma Aldrich, Cat No. R8875) were prepared in the unsupplemented medium to provide a final concentration of 1µM of oligomycin A, 0.5µM of FCCP and 1µM of antimycin A/rotenone in each well. The individual drugs were pipetted into each port of the hydrated plate (Antimycin A and rotenone were pipetted into the same port), in a sequential order (oligomycin A, FCCP, antimycin A/rotenone), and the hydrated plate now containing the drugs was placed in the Seahorse machine to be calibrated.

Once the cells and Seahorse machine were ready, the cells were placed in the machine, for the following programme; mix the medium for three minutes, wait for two minutes, measure the OCR and ECAR over three minutes, inject from the first port containing oligomycin, this protocol was repeated for the three ports containing drugs. Drugs were injected in the order oligomycin, FCCP and antimycin A/Rotenone.

To normalize the values provided by the Seahorse software to cell number, the cells were fixed using 4% paraformaldehyde (PFA) (Sigma Aldrich, Cat No. 158127) and stained with DAPI (Sigma Aldrich, Cat No.D8417); briefly, the cells were washed carefully with PBS, 450µl of 4% PFA was added to each well and the cells were incubated at room temperature for 15 minutes, before being washed again with PBS. The PBS was then removed and the cells were incubated in the dark for 15 minutes with 6µg/ml of DAPI dissolved in PBS. The cells were washed twice with PBS and kept hydrated in 50µl of PBS whilst being photographed using an inverted fluorescence microscope. Open CFU software (235) was used to quantify the number of cells in each well.

Cell density titration experiments were carried out to determine the optimal number of cells for the Seahorse analysis

8.7. Statistical analysis

Data are shown as ± standard deviation from the mean, represented as error bars. Differences from the mean were calculated using the Students' unpaired t-test for comparisons between two samples, and ANOVA for comparisons between more than two samples using the GraphPad Prism software (236) to determine if there were any significant differences in expression. If the p-value was less than 0.05, then further posthoc t-test were conducted. Differences were considered significant if the p value was ≤ 0.05 represented by *, ** represents p values of ≤ 0.01 and *** represents p values of ≤ 0.01.

9. Results

9.1. WT and DGCR8$^{-/-}$ mESCs

All experiments in this chapter consisted of three biological replicates (three separate passages of each cell type). Each biological replicate was repeated twice, i.e. two flasks of the same cell type were plated and harvested at the same time therefore there were two technical replicates per passage. N=3 represents the number of biological replicates.

9.1.1 Characterisation

Initially, the DGCR8$^{-/-}$ mESCs were grown on a MEF feeder layer and expanded onto gelatin-coated plates, whereas the WT mESCs were grown directly on gelatin-coated plates. Therefore it was vital to characterise both cell types to ensure that they remained pluripotent.

DGCR8$^{-/-}$ and WT cells were grown for 5 days, fixed and stained for alkaline phosphatase. The number of cells that stained positively for alkaline phosphatase was an indicator of the number of pluripotent/stem cells in culture (Figure 13).

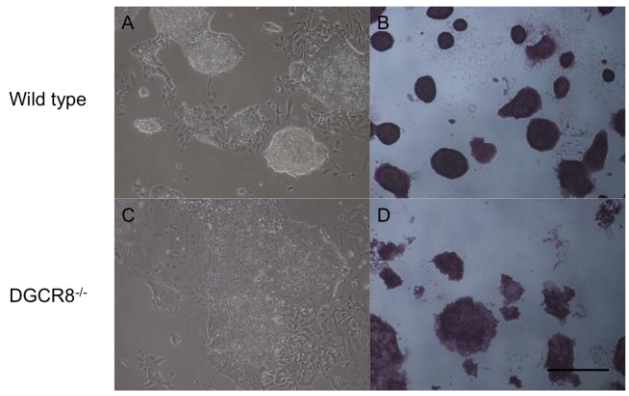


Figure 13 - Alkaline Phosphatase staining. The top row of images (A, B) are WT mESCs and the bottom row of images (C, D) are DGCR8$^{-/-}$ mESCs. Line represents 200µm

Although after 5 days in culture both cell types had begun to differentiate around the outer edges of their colonies, (Figure 13) they displayed 90-95% positive staining for alkaline phosphatase. The WT formed rounder, more compact colonies than the DGCR8$^{-/-}$ mESCs.

PCRs and immunoblotting were used to confirm the expression of the three pluripotency markers, Oct4, Nanog and Sox2, in wild type and DGCR8$^{-/-}$ mESCs (Figure 14).

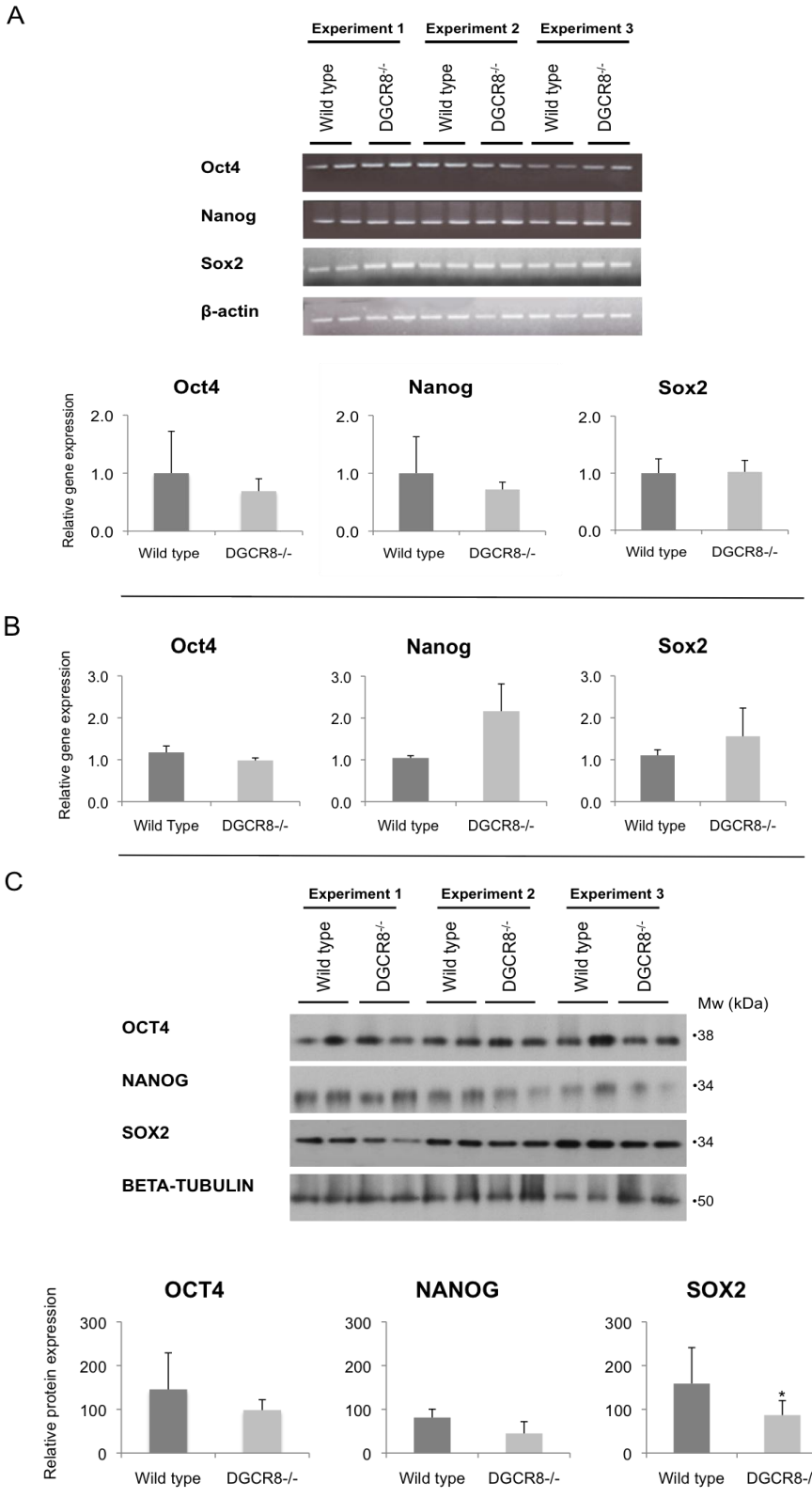


Figure 14 – Expression of pluripotency markers. A) Conventional PCR and densitometry normalised to loading control β-actin, B) QPCR normalised to positive control, β-actin, C) Immunoblot and densitometry normalised to loading control β-tubulin. p value <0.05 (t-test), n=3

The pluripotency markers Oct4, Nanog and Sox2 showed similar expression levels in both the DGCR8$^{-/-}$ and WT cells using both the conventional (Figure 14a) and quantitative PCR methods (Figure 14b). In accordance with the PCR data (Figures 14a, b) the two pluripotency markers Oct4 and Nanog were expressed at similar levels in both cell types at the protein level. Unexpectedly, Sox2 showed differential expression at the protein level (Figure 14c).

DGCR8 is a key protein in the miRNA biogenesis pathway, therefore its loss results in the loss of miRNAs (Figure 15).

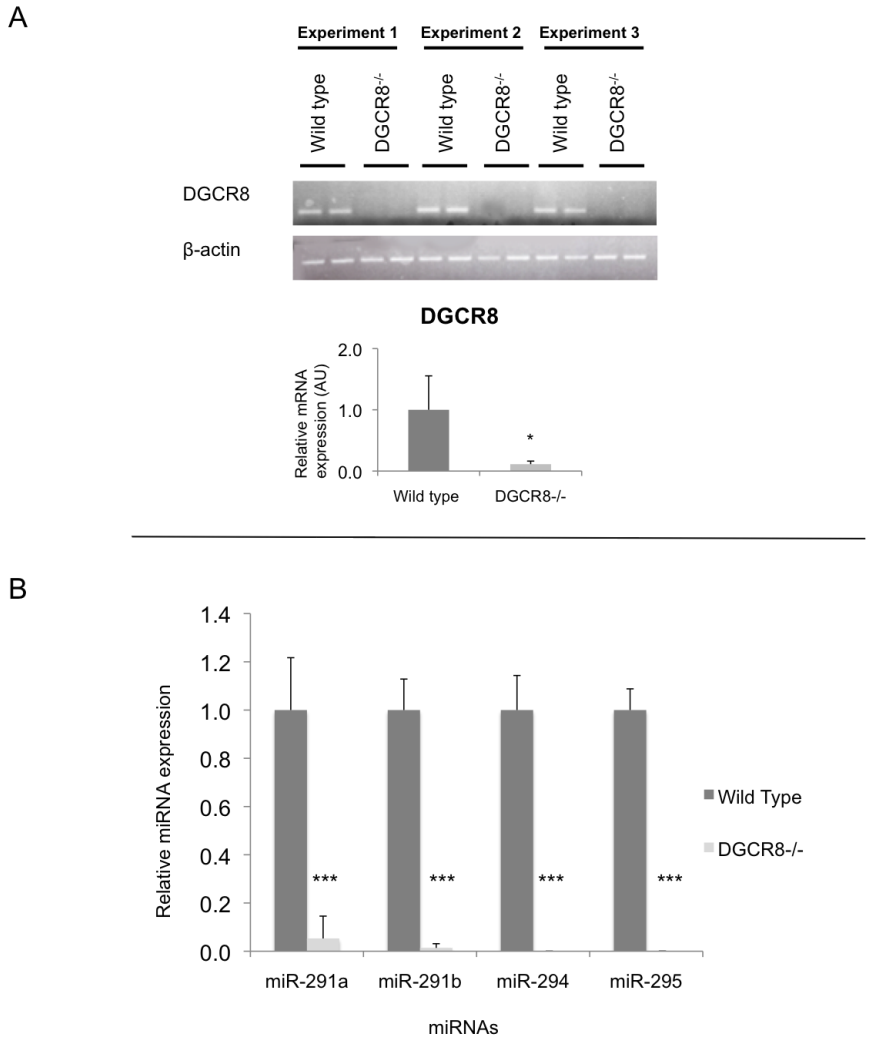


Figure 15 – Genotype of WT and DGCR8$^{-/-}$ mESCs and expression levels of miRNAs involved in the embryonic stem cell cycle. A) DGCR8 mRNA expression levels were normalized to β-actin. Wild type expression levels were set to 1, B) miRNA expression levels were normalized to U6 (n=3 independent experiments). MiRNA expression in wild type cells was set as 1. P(t-test) < 0.05

DGCR8$^{-/-}$ mESCs are known to express significantly higher levels of cell cycle inhibitors compared to WT cells (6), therefore we also evaluated the expression of the cell cycle inhibitors (Figure 16).

DGCR8 was undetectable in the knockout cells (Figure 15a) and this resulted in a significantly lower expression of the four canonical miRNAs (miR-291a, miR-291b, miR-294 and miR-295) in the DGCR8$^{-/-}$ mESCs compared to the WT mESCs (Figure 15b).

Further experiments were conducted on the cells to confirm previous findings, which showed that the DGCR8$^{-/-}$ mESCs displayed a slower proliferation rate compared to the WT cells (3,6). The specific loss of the miRNAs tested in Figure15, is consistent with the derepression of their targets, such as Cdkn1a, Rbl2 and Lats2 (Figure 16).

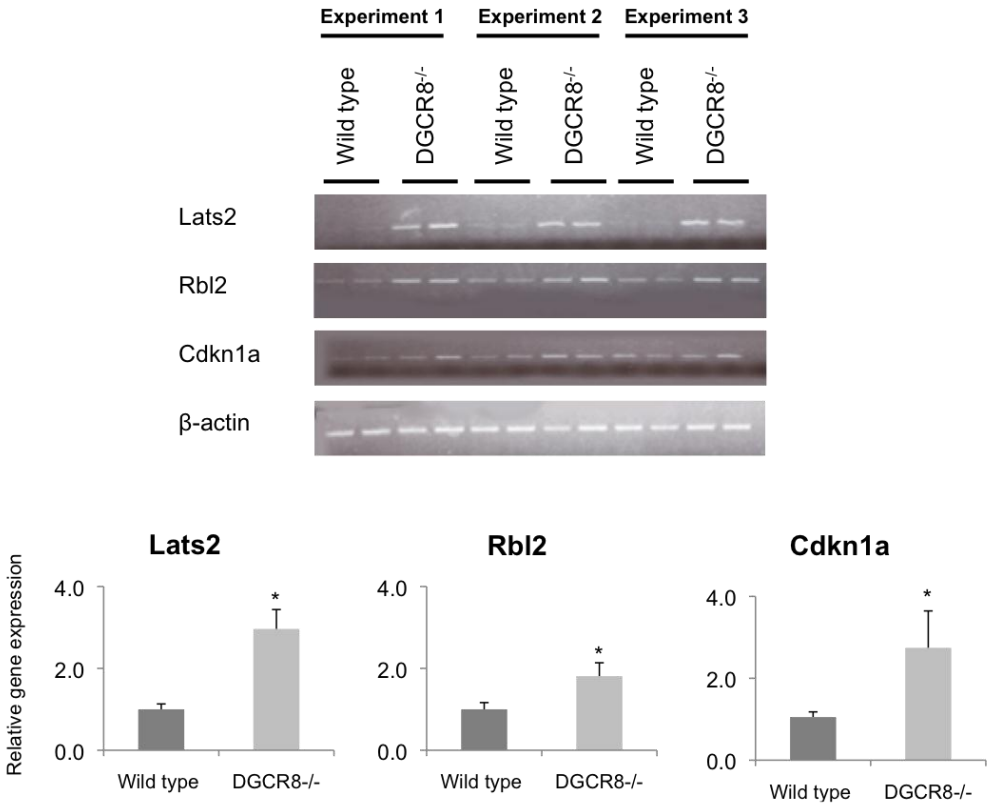


Figure 16 – mRNA expression levels of cell cycle inhibitors. Conventional PCR and densitometry normalised to loading control. QPCR was conducted on Cdkn1a and normalized to positive control. Wild type expression levels are set to 1. P(t-test) < 0.05,n=3

Lats2, Rbl2 and Cdkn1a are highly expressed inhibitors of the G_1/S phase transition in WT mESCs, however the results suggest that only Lats2 and Rbl2 have significant differential expression in the two cell types (Figure 16). However using qPCR, which is a more sensitive and accurate gene quantification method, Cdkn1a showed a significantly higher expression in the DGCR8$^{-/-}$ compared to the WT mESCs, which is in agreement with previous studies (3).

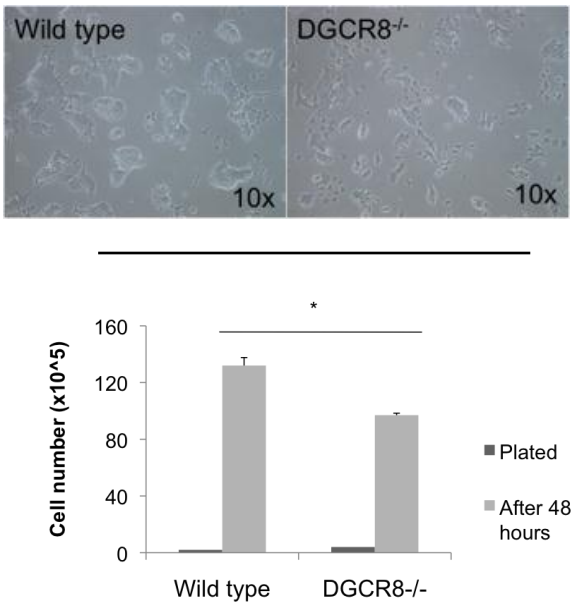


Figure 17 – Cell numbers. Analysis was conducted after 48 hours in culture. n=4 (two technical replicates in two independent experiments). T-test was conducted between WT and DGCR8$^{-/-}$ mESCs, P(t-test) <0.05

As reported previously, the DGCR8$^{-/-}$ cells were displaying the expected phenotype of a slow proliferation rate compared to WT cells (6). Figure 17 indicates that although half the number of WT mESCs compared to DGCR8$^{-/-}$ mESCs were plated, after 48 hours of growth under identical conditions, there were significantly more WT mESCs compared to DGCR8$^{-/-}$ mESCs.

To ensure that the differences in cell number were solely due to proliferation defects as opposed to cell death via apoptosis, a cell death detection assay was conducted, which quantified the release of nucleosomes released by the cells into the growth medium (Figure 18).

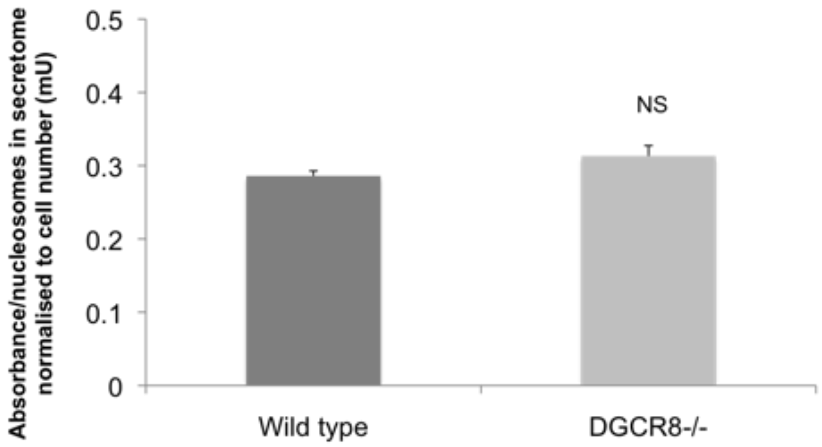


Figure 18 - Cell death detection assay. The number of nucleosomes released into the medium by the WT and DGCR8$^{-/-}$ mESCs, after 48 hours of growth was taken as an indication of the amount of cell death. Values were normalised to cell number after 48 hours. n=4 (two technical replicates in two independent experiments), error bars represent standard deviation. P(t-test) < 0.05. NS denotes not significant.

The cell death assay (Figure 18) demonstrates that the loss of canonical miRNAs does not result in increased apoptosis.

9.1.2 DIGE

2-D DIGE provides a quantitative method of comparing proteomes, by differentially labelling the samples with different fluorescent dyes (Cy3 (red) and Cy5 (green)), and separating the proteins from two samples plus a pooled standard (Cy2) on the same two-dimensional gel which allowed for reliable quantification of proteins between gels. 2-D DIGE analysis was carried out on DGCR8$^{-/-}$ and WT whole cell protein lysates, to compare differential protein expression (Figure 19). Proteins are separated by molecular weight (Mw) and isoelectric point (pI).

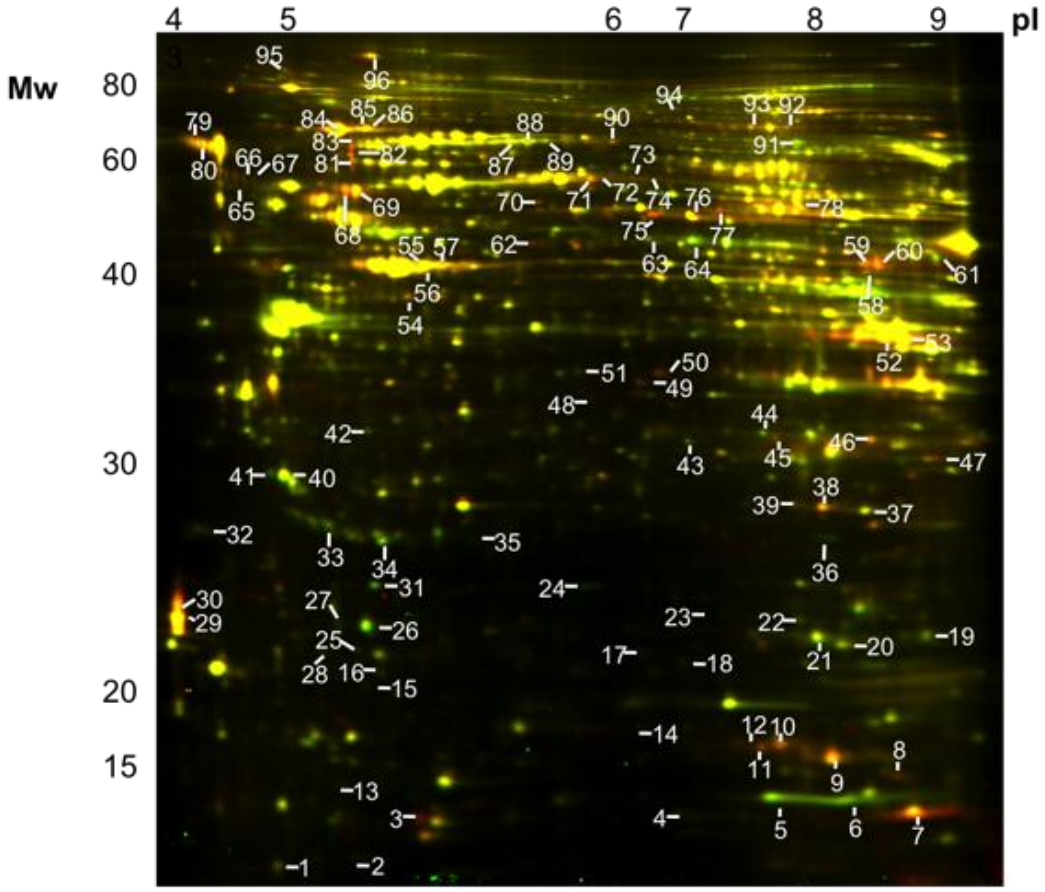


Figure 19 - Representative DIGE gel image. DCGR8$^{-/-}$ cells were labelled with Cy3 (red) and WT mESCs with Cy5 (green). A sample pool was labelled with Cy2 as an internal standard. Proteins were separated by molecular weight (Mw) and isoelectric point (pI). Spots that are orange or green represent proteins that are differentially expressed, yellow spots represent proteins that are expressed to the same level in both samples. Spots were excised if they showed at least a 1.5 fold change between cell types with a p-value of <0.05 (t-test), as determined by the Decyder software (GE Healthcare). Mw=mass of protein, pI= isoelectric point. Protein names are listed in *Table 7*.

96 spots that were differentially expressed in the two cell types were selected for LC-MS/MS analysis based on the parameters, of having at least a 1.5 fold change between the cell types at a significance level of $p<0.05$ (Figure 19). Of the 96 spots, 84 unique proteins were identified and from these 48 were upregulated in the DGCR8$^{-/-}$ cells compared to the WT cells, as would be expected for putative direct targets of the canonical miRNAs (Table 7).

Results from an Ingenuity pathway analysis software (237) to screen for any common pathways between proteins or any networks that the proteins may be

involved in are summarized in Figure 20. The two largest groups of proteins that showed differential expression between the two cell types were those involved in energy metabolism, and those with and related to chaperone function (Table 7).

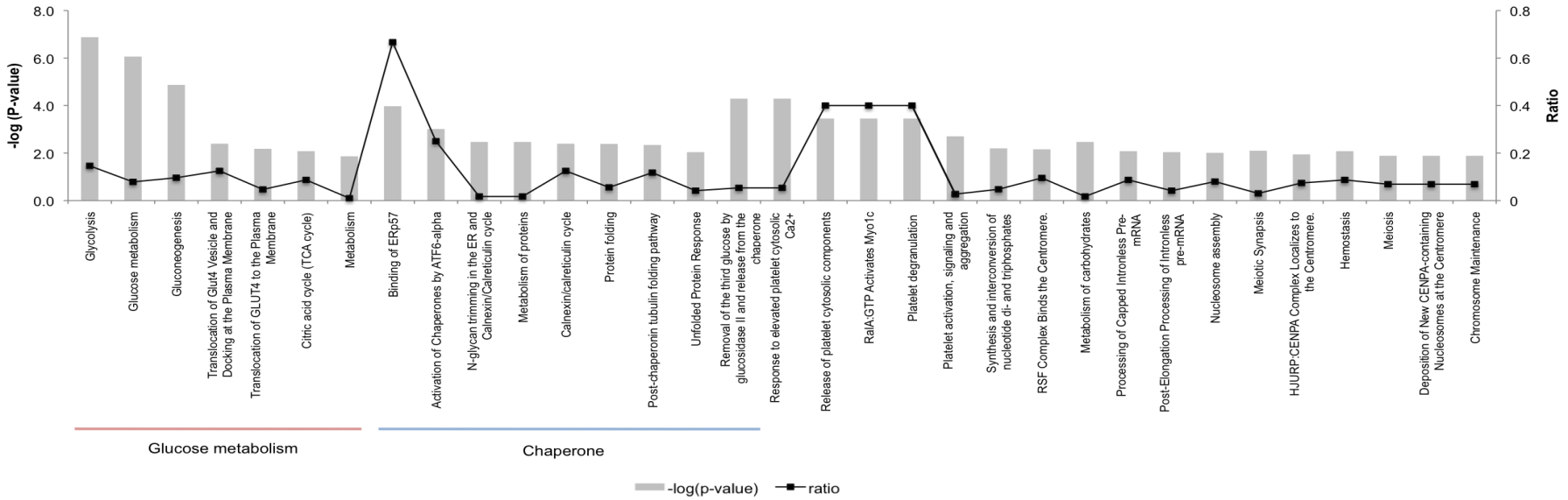


Figure 20 – Pathway analysis. Pathways affected by the loss of DGCR8 in mESCs as determined by REACTOME within the bioprofiling software (238,239) Pathways are ordered by level of significance, as determined by a hypergeometric test, which tests the 'probability of seeing n or more genes by chance', only proteins with a p-value less than 0.01 are shown. The ratio represents the number of proteins found in the DGCR8$^{-/-}$ vs. WT dataset that are present in the pathway divided by the total number of proteins in that pathway.

The biggest group of pathways affected by the loss of DGCR8 in mESCs were those with a chaperone function, however the top pathways are involved in glucose metabolism (Figure 20).

9.1.3 Validation of DIGE analysis (and associated proteins) of wild type and DGCR8$^{-/-}$ mESCs

In order to validate the results from the DIGE analysis (Figure 19 and Table 7) as well as the accompanying common pathway analysis (Figure 20); Western blots (Figure 21) and qPCRs (Figure 22) were conducted not only analysing the proteins identified in the DIGE study but those in associated pathways, such as glycolysis and the TCA cycle.

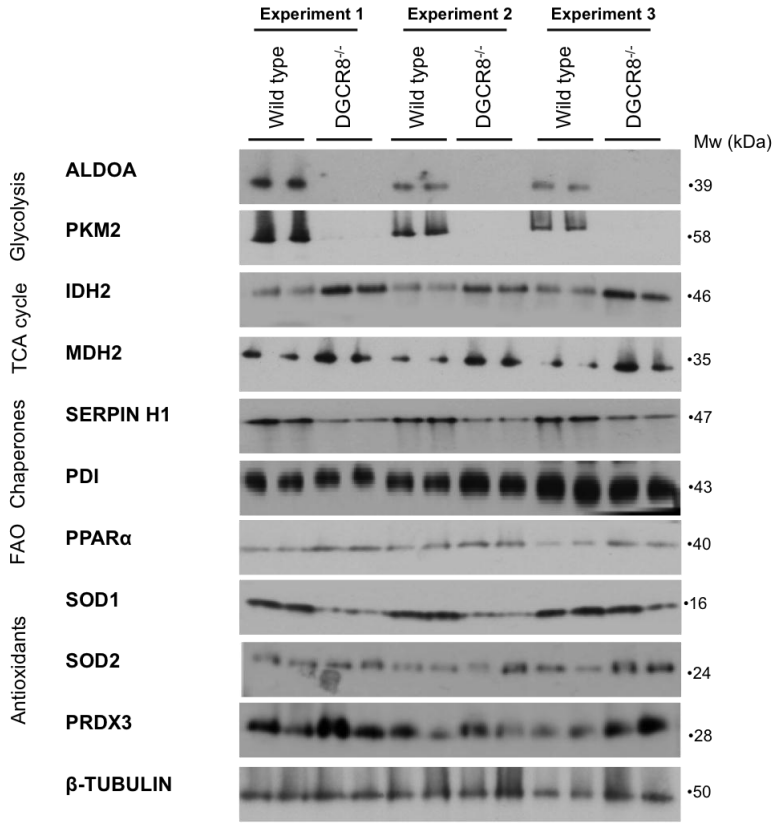


Figure 21 - Protein expression profile. Immunoblotting was conducted on 3 consecutive passages of WT and DGCR8$^{-/-}$ mESCs, labelled as experiment 1, 2 and 3. Protein targets were chosen based on their differential expression between the two cell types (Table 7) and their relationship with identified proteins. β-tubulin was used as a loading control. *Abbreviations- ALDOA: Fructose-bisphosphate Aldolase A, PKM2: Pyruvate kinase muscle isoform 2, IDH2: Isocitrate dehydrogenase 2, MDH2: Malate dehydrogenase 2 PDI: Protein disulphide isomerase, , PPARα: Peroxisome proliferator-activated receptor alpha, SOD1/2- Superoxide dismutase 1/2, PRDX3: Peroxiredoxin 3 FAO: Fatty acid oxidation*

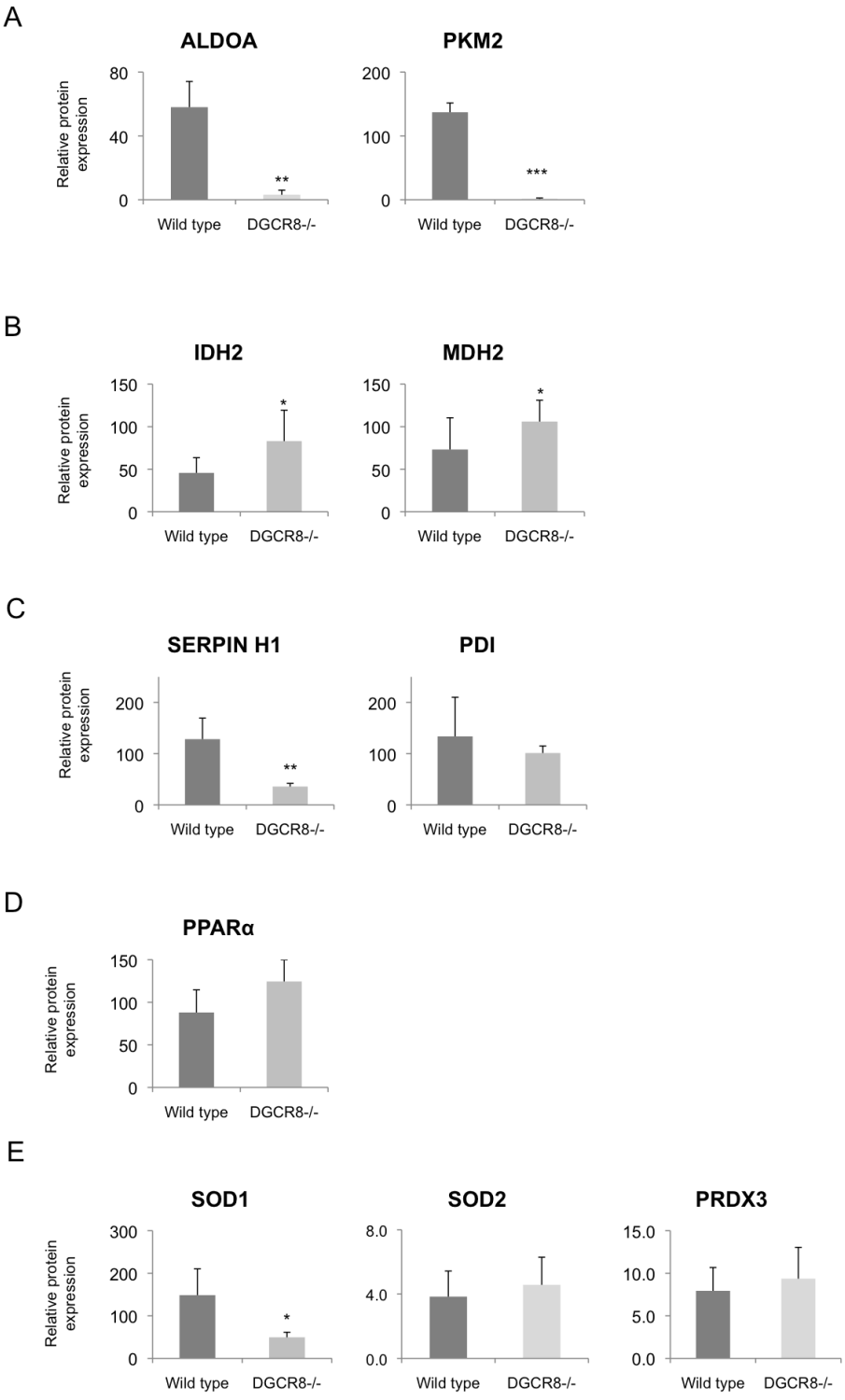


Figure 22 – Densitometry. A) Glycolytic enzymes, B) TCA enzymes, C) Chaperone proteins, D) Fatty acid metabolism, E) Antioxidants, n=3, P(t-test) <0.05. *Abbreviations- ALDOA: Fructose Bisphosphate Aldolase A, PKM2: Pyruvate Kinase muscle Isoform 2, PPARα: Peroxisome proliferator-activated receptor alpha, SOD1/2: Superoxide dismutase 1/2, PRDX3: Peroxiredoxin 3.*

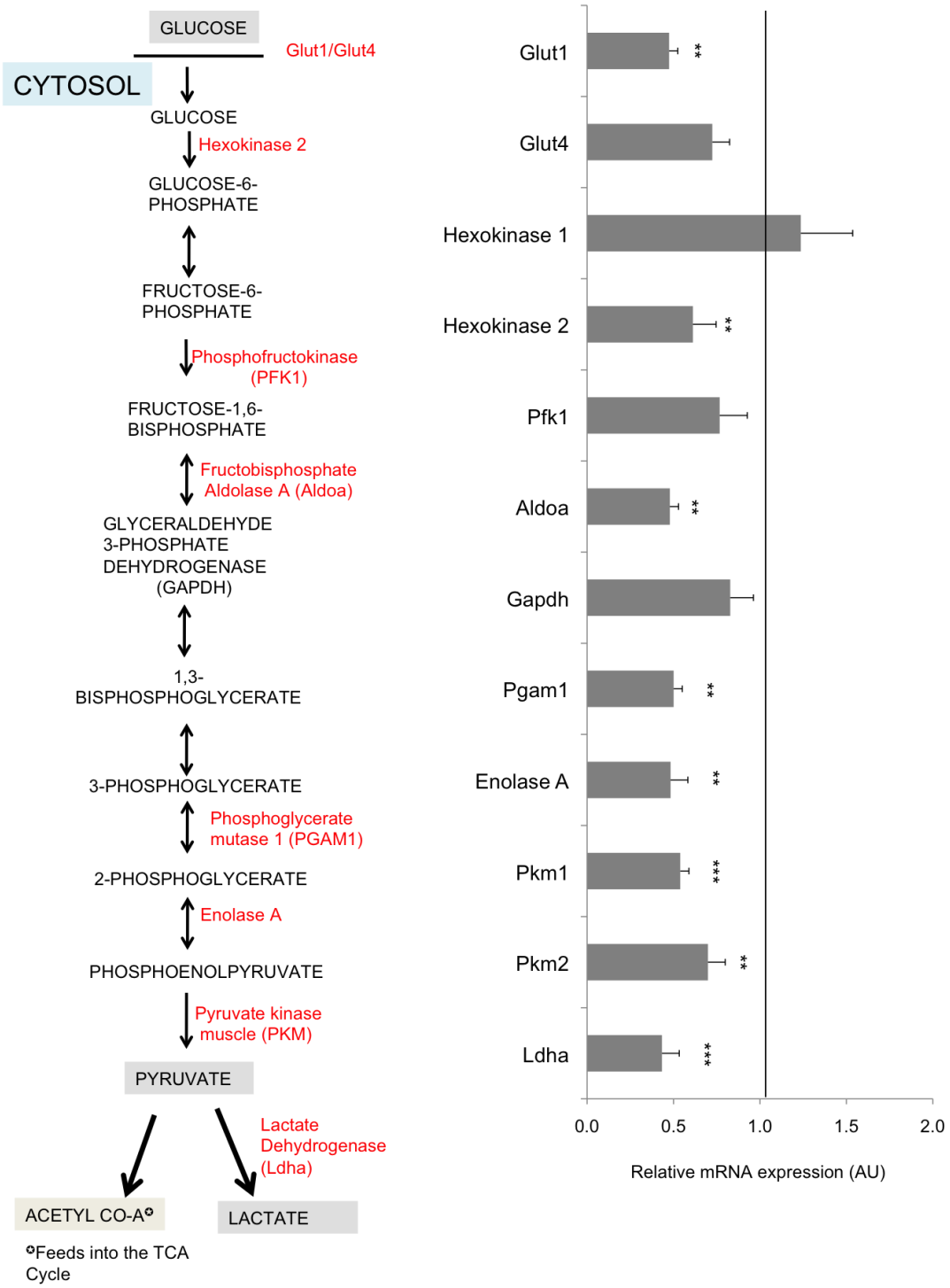
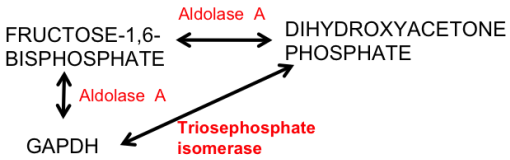
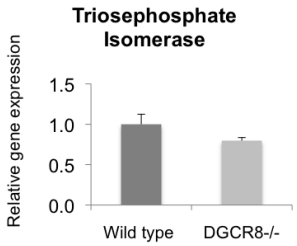


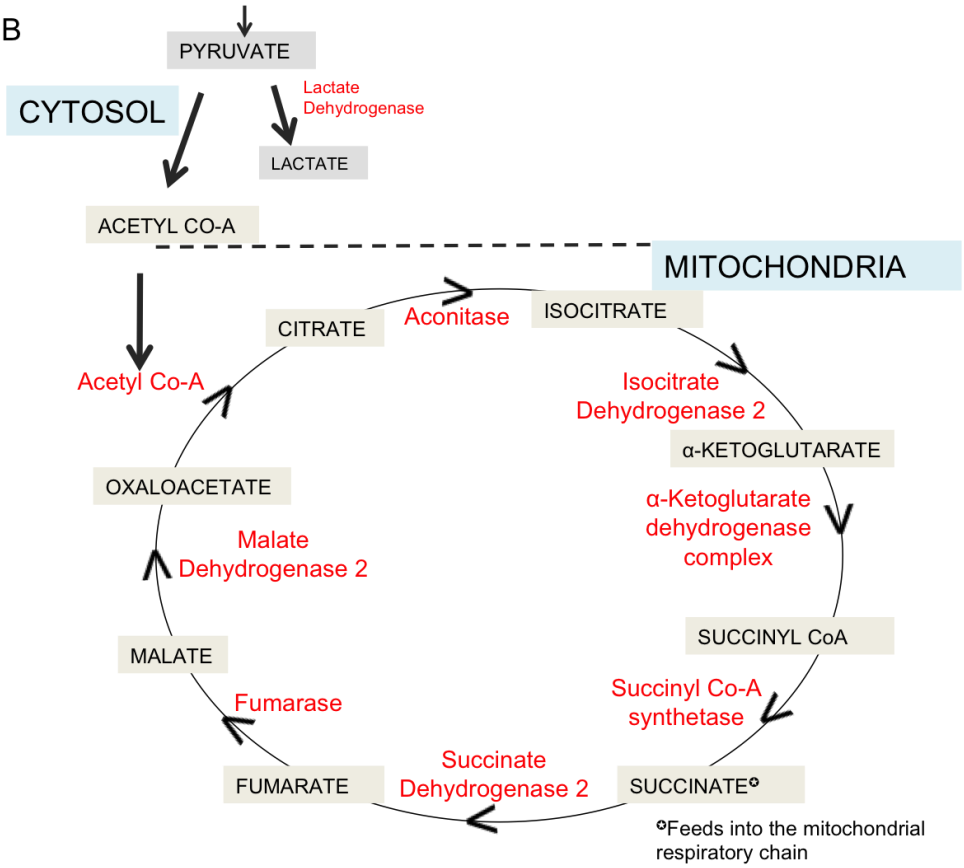
Figure 23 - Quantitative PCR profile. Expression of glycolytic genes in WT and DGCR8$^{-/-}$ mESCs. Expression levels were normalised to the loading control β-actin. Expression in WT mESCs was set as one, represented by the horizontal line. n=3, P(t-test) < 0.05

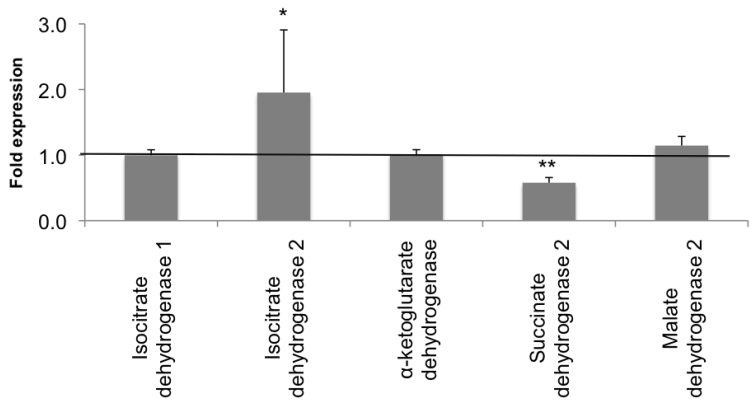
A


FRUCTOSE-1,6-BISPHOSPHATE
Aldolase A
DIHYDROXYACETONE PHOSPHATE
Aldolase A
GAPDH
Triosephosphate isomerase


Triosephosphate Isomerase
Relative gene expression
1.5
1.0
0.5
0.0
Wild type
DGCR8-/-

B


PYRUVATE
CYTOSOL
Lactate Dehydrogenase
LACTATE
ACETYL CO-A
MITOCHONDRIA
CITRATE
Aconitase
ISOCITRATE
Acetyl Co-A
Isocitrate Dehydrogenase 2
α-KETOGLUTARATE
OXALOACETATE
α-Ketoglutarate dehydrogenase complex
Malate Dehydrogenase 2
SUCCINYL CoA
MALATE
Fumarase
Succinyl Co-A synthetase
FUMARATE
Succinate Dehydrogenase 2
SUCCINATE✪
✪Feeds into the mitochondrial respiratory chain


Fold expression
3.0
2.0
1.0
0.0
*
**
Isocitrate dehydrogenase 1
Isocitrate dehydrogenase 2
α-ketoglutarate dehydrogenase
Succinate dehydrogenase 2
Malate dehydrogenase 2

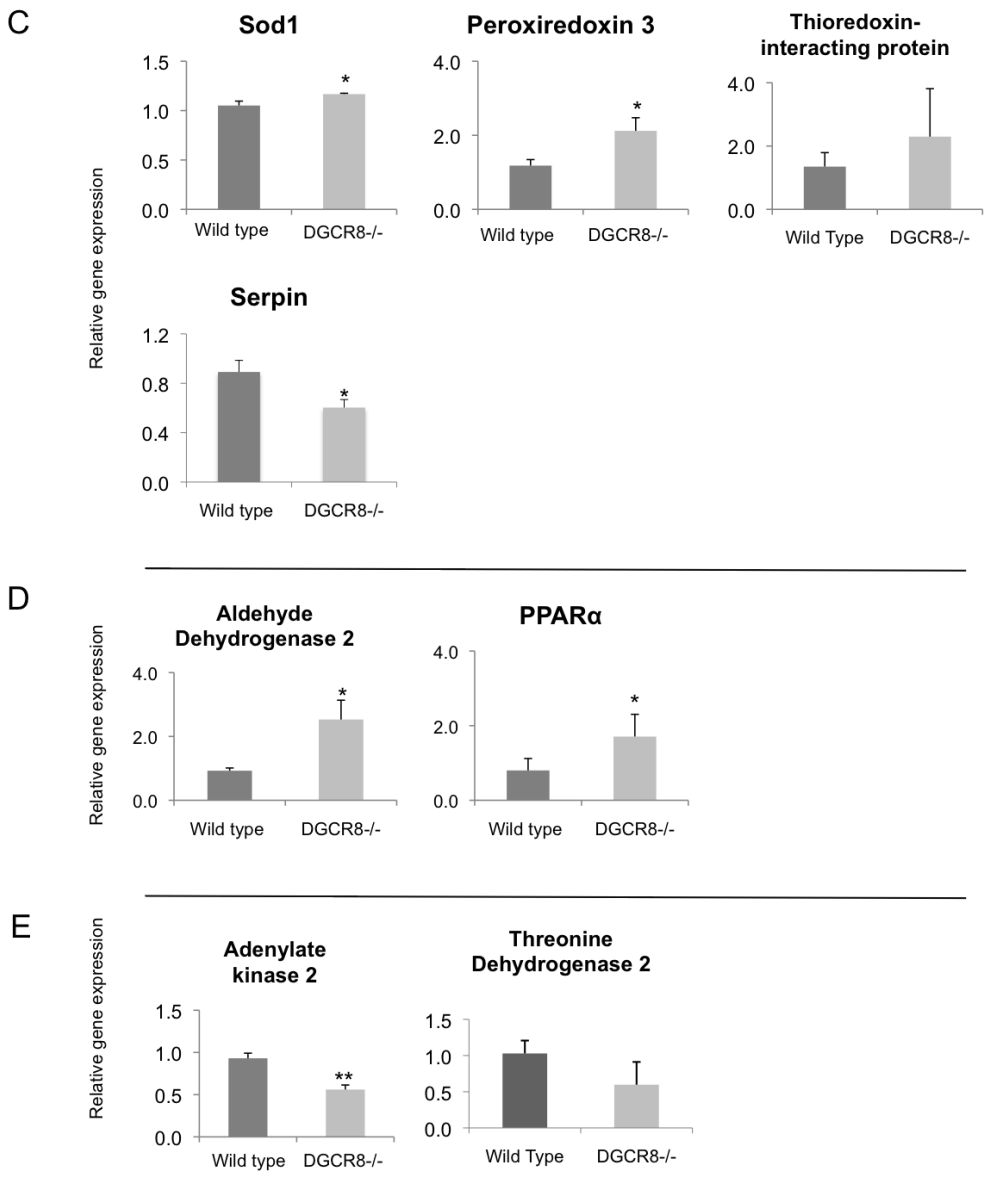


Figure 24 - Quantitative PCR profile (2). Gene expression in WT and DGCR8$^{-/-}$ mESCs**.** A) Associated with glycolysis, B) TCA cycle, C) Antioxidant and chaperone, D) Fatty acid metabolism, E) Associated with energy metabolism. Expression levels were normalised to the loading control β-actin. Expression in wild type mESCs was set as one. n=3,P(t-test) < 0.05

The RNA and protein analysis revealed differences in the expression of glucose metabolism enzymes between the DGCR8$^{-/-}$ and WT mESCs, with the DGCR8$^{-/-}$ mESCs displaying lower expression of glycolytic enzymes (Figures 21, 23) and a reciprocal higher expression of TCA cycle enzymes such as Malate Dehydrogenase 2 and Isocitrate Dehydrogenase 2 (Figures 21, 24b), the opposite trend was observed in the WT cells.

In agreement with the pathway analysis presented in Figure 20, eight out of the twelve enzymes involved directly in glycolysis displayed lower expression in the DGCR8$^{-/-}$ mESCs compared to the WT (Figure 23).

Other proteins including those with antioxidant function; Sod1 and a chaperone protein; Serpin H1, were also identified as having differential expression between the two cell types and were validated at both the protein and RNA levels (Figures 22, 24).

9.1.4 Metabolic analysis

As glycolytic and TCA cycle enzymes were differentially expressed between the two cell types, the Seahorse XF^{e} analyzer was used to further study glucose metabolism in the WT and DGCR8$^{-/-}$ mESCs (Figure 25).

The Seahorse XF^{e} analyzer, allows for the simultaneous measurement of the amount of mitochondrial respiration that occurs within cells/tissues through a parameter termed the Oxygen Consumption Rate (OCR) and the number of protons produced by the cells/tissues via parameter termed the Extracellular Acidification Rate (ECAR). The OCR and ECAR provide readouts for the amount of oxidative phosphorylation and glycolysis occurring in cells, respectively. The analyzer not only allows for the simultaneous measurement of the ECAR and OCR over a defined period of time, but also for chemical modulators to be directly added to the cells, permitting the study of mitochondrial function and glycolytic capacity of the cells studied (Figure 25).

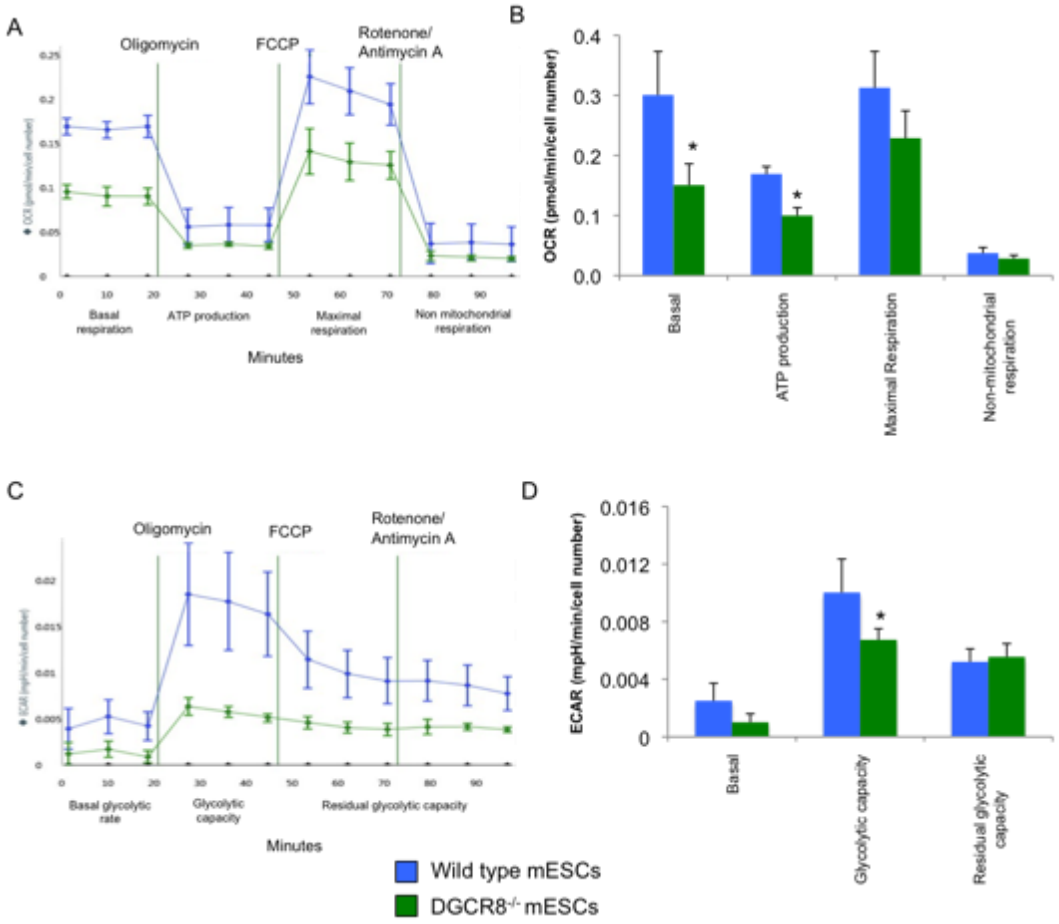


Figure 25 – Seahorse analysis of WT and DGCR8$^{-/-}$ mESCs. A) Representative OCR (Oxygen Consumption Rate) of WT and DGCR8$^{-/-}$ mESCs in the presence of three mitochondrial modulators; Oligomycin (ATP synthase inhibitor), FCCP (mitochondrial uncoupler) and Rotenone/Antimycin A (Complex I/Complex IV inhibitor), B) Summary of OCR measurements, C) Representative ECAR (Extracellular Acidification Rate) of WT and DGCR8$^{-/-}$ mESCs in the presence of three mitochondrial modulators; Oligomycin (ATP synthase inhibitor), FCCP (mitochondrial uncoupler) and Rotenone/Antimycin A (Complex I/Complex IV inhibitor), D) Summary of OCR measurements. n=3, P(t-test) <0.05

Although both cell types, had a similar basal and residual glycolytic rate, the WT cells had a significantly higher glycolytic capacity (Figure 25c, d), which is line with their higher mRNA and protein expression of glycolysis-associated enzymes (Figures 21, 23). At baseline, the DGCR8$^{-/-}$ cells had a significantly lower OCR

(Figure 25b). This was also the case upon the addition of Oligomycin A, an ATP synthase inhibitor.

The inhibition of ATP synthase by Oligomycin A led to a fall in ATP production and a reciprocal increase in the ECAR, as the cells attempted to compensate for the shortfall in ATP production, indicative of the higher glycolytic capacity of the WT cells compared to the DGCR8$^{-/-}$ cells (Figure 25d). The maximal respiration rate (Figure 25b) was calculated using the OCR values upon the addition of FCCP minus the non-mitochondrial respiration (values collected upon the addition of Rotenone and Antimycin A), the WT cells tended to show a higher maximal respiration rate overall, however the difference was not significant (Figure 25c).

The metabolite concentrations of WT and DGCR8$^{-/-}$ mESCs were studied (Figure 26) in their conditioned medium. The values of the metabolites quantified using H^1 NMR spectroscopy are summarized in Table 8.

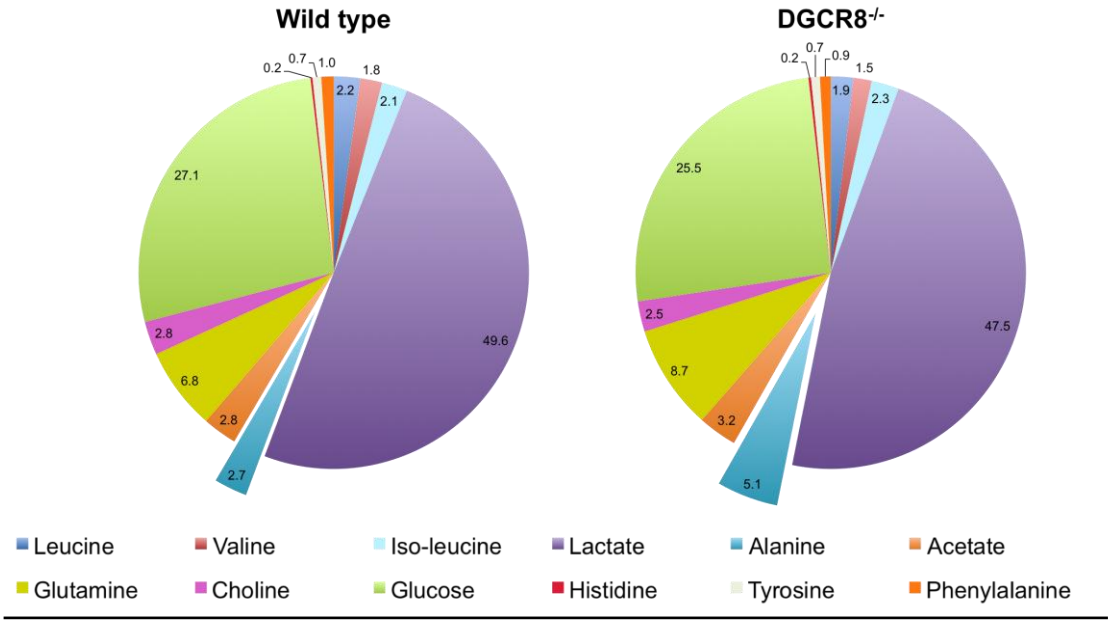


Figure 26- Metabolites in cell culture supernatants. Values represent percentages and have been normalised to cell numbers after 48 hours of growth. P (t-test) < 0.05

The metabolomic studies revealed that after 48 hours of growth, in both cases glucose was taken up at a higher proportion than any other metabolite and

lactate was excreted the most (Table 8). Alanine, a transamination product of pyruvate and a precursor for glutaminolysis, was the only metabolite that showed a significant difference between the two cell types (Figure 26 and Table 8).

9.2 Transfection studies

To further explore the role of miRNAs we transfected the DGCR8$^{-/-}$ mESCs with four miRNAs from the 295 cluster; miR-291a, miR-291b, miR-294 and miR-295, and miR-302d from the 302 cluster. The efficiency of transfection was confirmed via qPCR (Table 10).

9.2.1 Characterisation after transfection

PCRs and immunoblotting were used to confirm the expression of the three pluripotency markers, Oct4, Nanog and Sox2, in WT and DGCR8$^{-/-}$ mESCs transfected with miRNAs from the 295 cluster and miR-302d (Figure 27). Cdkn1a, a confirmed direct target of the transfected miRNAs, was also tested.

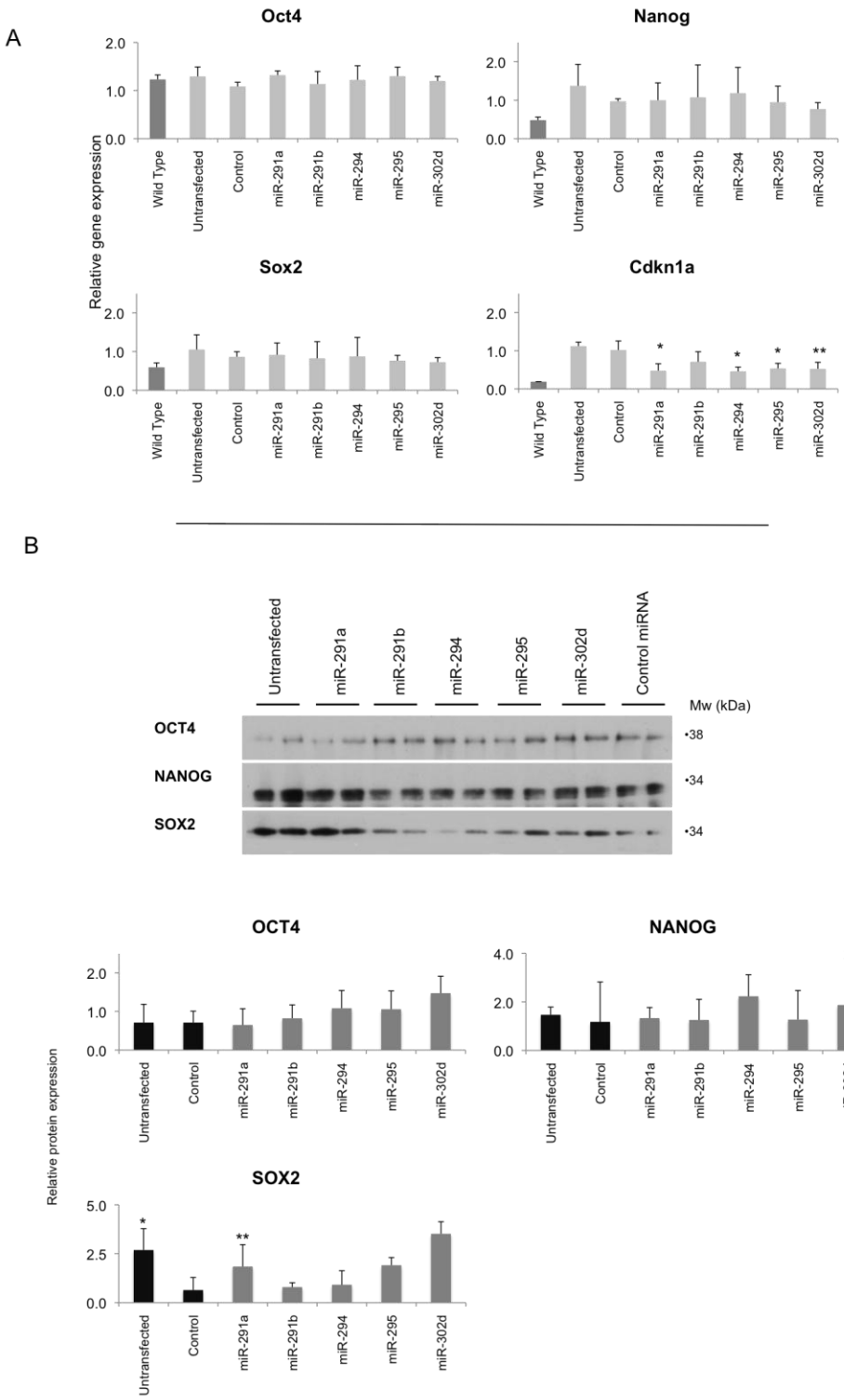


Figure 27- Expression of pluripotency markers and cell cycle inhibitor in DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs. Pluripotency (Oct4, Nanog and Sox2) and cell cycle inhibitor (cdkn1a) genes. Gene expression levels were normalised to the loading control β-actin, n=3, P(ANOVA)<0.05. *Abbreviation- cdkn1a-cyclin-dependent kinase inhibitor 1*

The pluripotency markers, Oct4, Nanog and Sox2 remained unaffected by the over expression of the ESCC miRNAs (Figure 27).

Cdkn1a (G_1/S phase cell cycle regulator) that has previously been identified as a target of the ESCC miRNAs (3), was used a positive control to ensure that the expression of the ESCC in the DGCR8$^{-/-}$ mESCs resulted in its downregulation (Figure 27).

The cells were also counted just before they were harvested to ensure that the transfection was not detrimental to cell viability (Figure 28b). Although on average the cell numbers before harvest were similar between transfections, individually, the transfection of miR-291b resulted in a significantly lower final cell number (Table 11), therefore it was eliminated from some of the further analysis.

On average after 48 hours of transfection with the ESCC miRNAs there was no significant difference in cell number (Figure 28b), this is supported by the photos of DGCR8$^{-/-}$ transfected with the ESCC miRNAs (Figure 28a). The transfected mESCs looked similar except for the mESCs transfected with miR-291b, which displayed less round colony formation and more elongated cells around the edges of the colonies (Figure 28a).

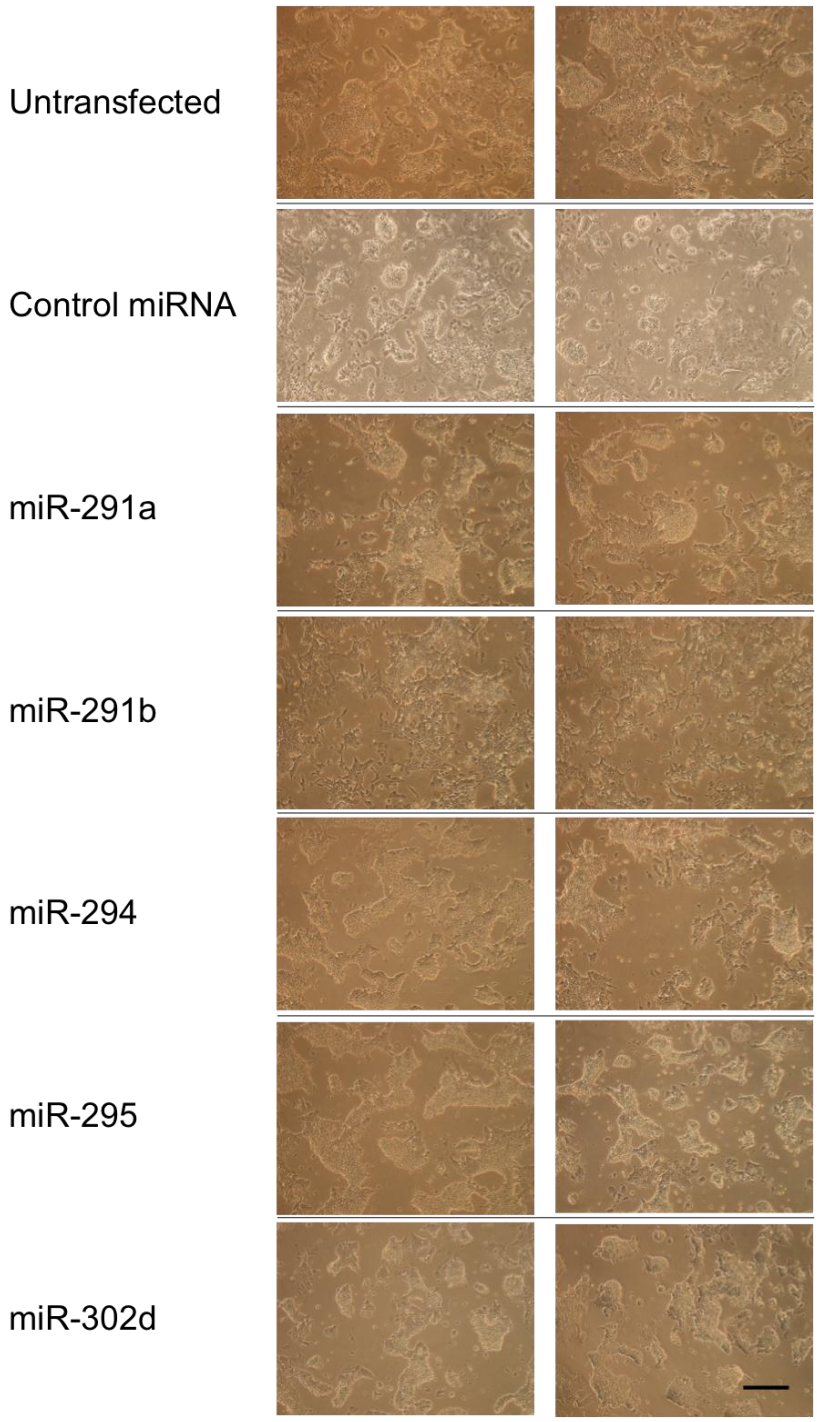


Figure 28a- Representative photos of transfected DGCR8$^{-/-}$ mESCs. Photos were taken 48 hours post transfection. Line represents 200µm

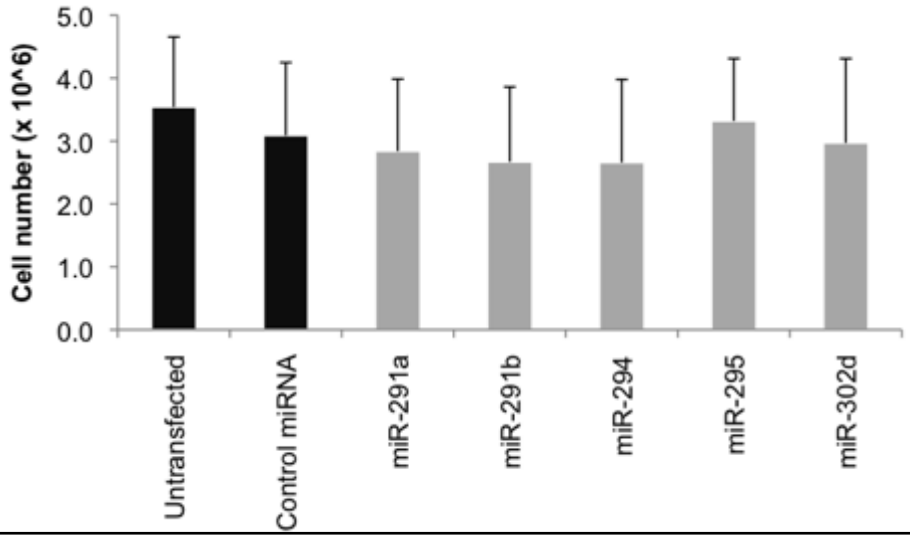


Figure 28b - Average cell numbers of transfected DGCR8$^{-/-}$ mESCs. Cells were individually transfected with the ESCC miRNAs and counted 48 hours post-transfection. Error bars represent standard deviation, P(ANOVA) < 0.05, none of the differences were found to be significant. n=6

9.2.2 Transfection DIGE

In order, to gauge the effects that the individual ESCC miRNAs have on the proteome of mESCs, comparative 2-D DIGE analysis was conducted on the transfected cell lysates(Figure 29). Protein changes upon transfection with individual ESCCs are listed in Table 9.

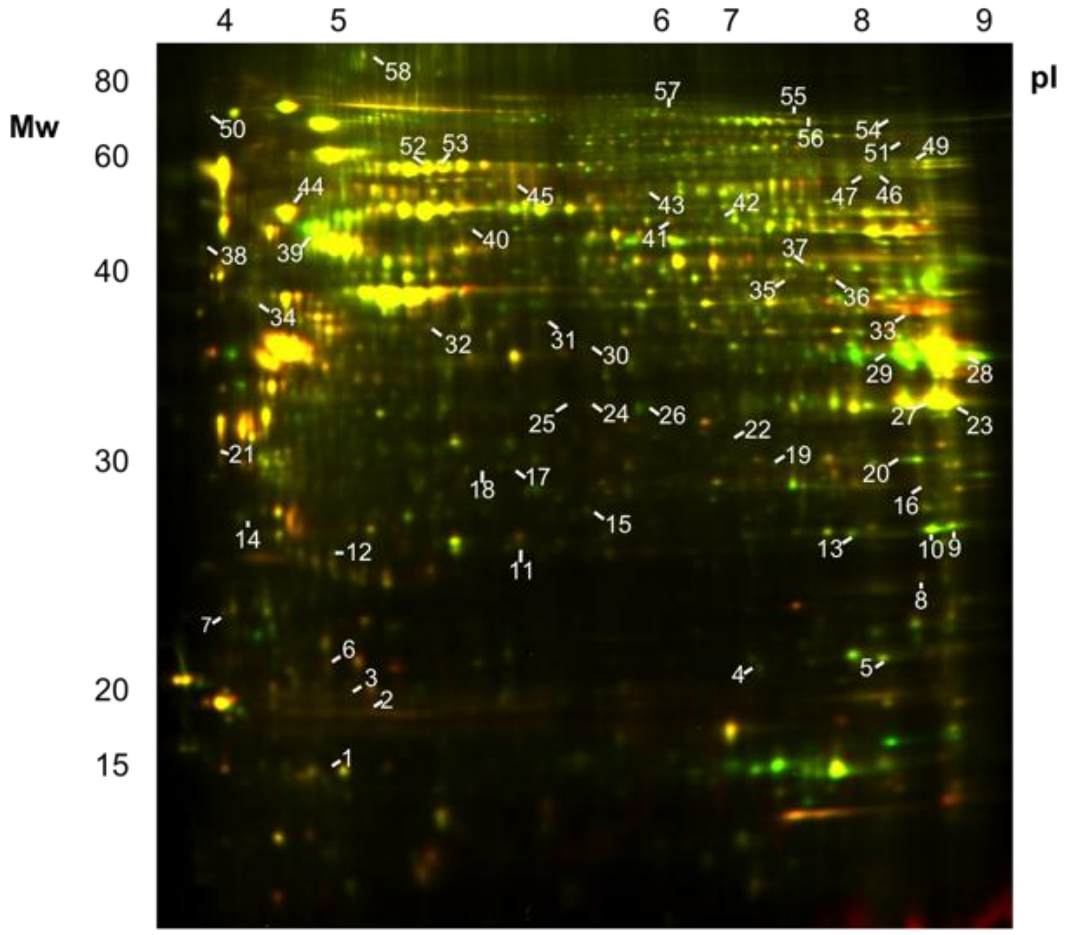


Figure 29 – Representative DIGE gel image comparing the differences in protein expression between DGCR8$^{-/-}$ mESCs transfected with miR-302d and the Control miRNA . The 58 numbered proteins represent proteins that show differences in protein expression above the threshold level between the DGCR8$^{-/-}$ transfected with each of the ESCC miRNAs and those transfected with the control miRNA. Cells transfected with the control miRNA were labelled with Cy3 (red) and DGCR8$^{-/-}$ cells transfected with miR-302d were labelled with Cy5 (green). Both transfected cells were labelled with Cy2 as an internal standard. Spots that are orange or green represent proteins that are differentially expressed, yellow spots represent proteins that are expressed to the same level in both samples. Spots were excised if they were within the parameters of a 1.2 fold change between cell types and a p value of <0.05, as determined by the Decyder software. Mw=mass of protein, pI= isoelectric point. Proteins are summarised in Table 9.

A representative DIGE gel of the DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs is shown in Figure 29. Proteins expressed in the DGCR8$^{-/-}$ mESCs transfected with the control miRNA were labelled with Cy3, and those DGCR8$^{-/-}$ mESCs transfected with miR-302d were labelled with Cy5.

58 of the differentially expressed spots were selected for LC-MS/MS analysis based on the parameters, of having at least a 1.2 fold change between the transfections at a significance level of $p<0.05$. The fold change threshold was set lower than in the initial DIGE comparison between the wild type and DGCR8$^{-/-}$ mESCs, to increase the power of the DIGE analysis and the chances of identifying real targets, as the five miRNAs have the same seed sequence and have many similar targets (3).

Of the 58 spots picked, 51 unique proteins were identified, of those 11 were identified in the first 2-D DIGE study as differentially expressed between the untransfected and DGCR8$^{-/-}$ mESCs (Table 14). Of these 11, 9 were upregulated in the DGCR8$^{-/-}$ mESCs compared to the WT mESCs, indicating that they may be targets of the ESCCs (Table 14).

Although 12 DIGE gels were run, encompassing all the individual transfections (two technical replicates and two biological replicates per transfection), only eleven gels were used for principal component analysis (Figure 30). This was due to the uneven running of one gel (containing a comparison between DGCR8$^{-/-}$ mESCs transfected individually, with a control miRNA and miR-302d). The technical artefact would have introduced greater experimental variability into the analysis.

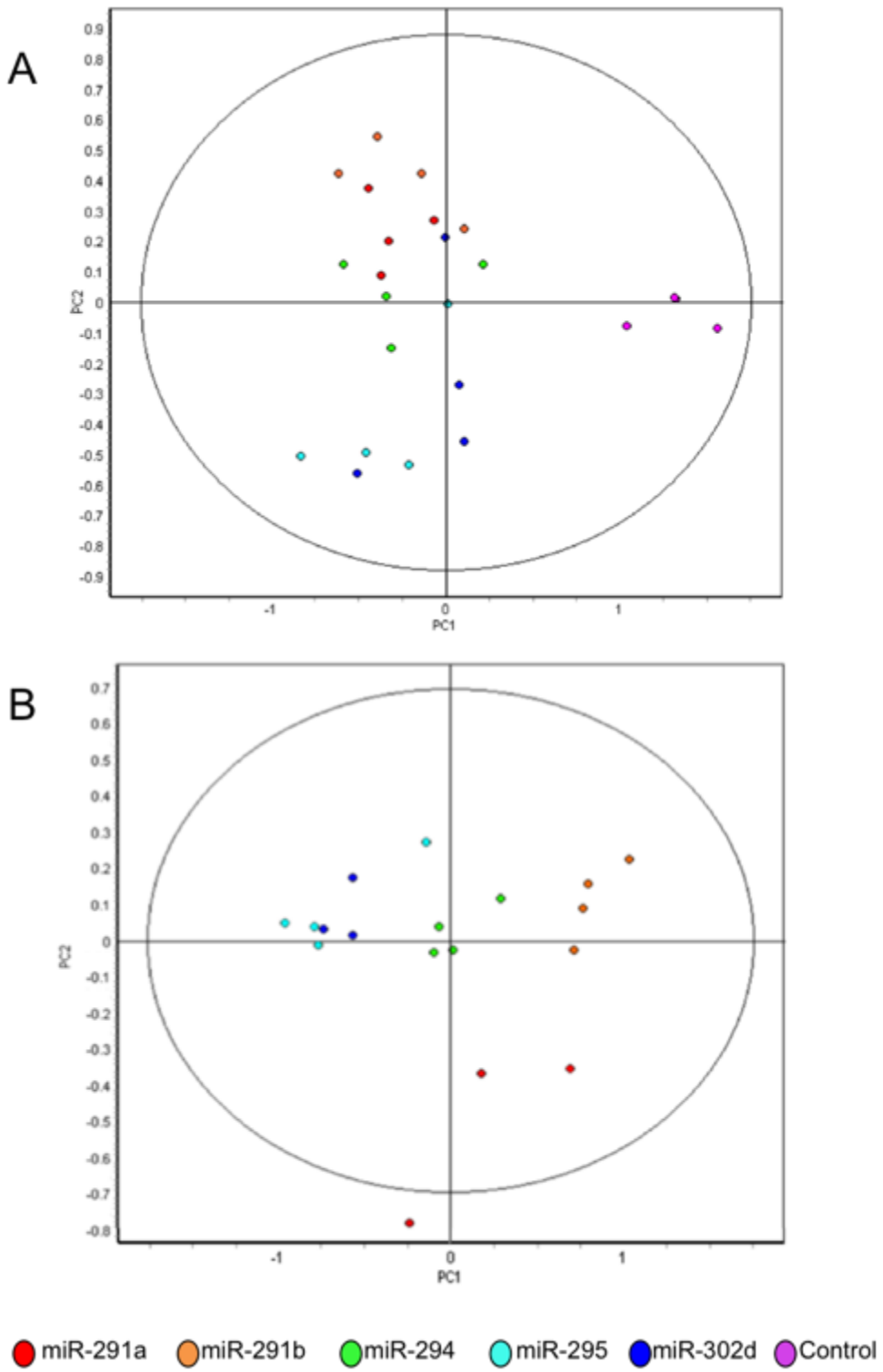


Figure 30 –Principal Component Analysis (PCA). PCA was conducted based on targets identified in the DIGE study as being affected by the ESCC miRNAs. a) PCA of the DGCR8$^{-/-}$ mESCs transfected with miR-291a, 291b, 294, 295, 302d and a control miRNA, b) PCA of the DGCR8$^{-/-}$ mESCs transfected with miR-291a, 291b, 294, 295 and 302d (n=3 or 4).

By conducting PCA on the DGCR8$^{-/-}$ mESCs transfected with the ESCC, it was established that the cells transfected with the control miRNA were distinct from the cells transfected with the ESCC miRNAs (Figure 30a) in terms of their effect on protein expression. For closer inspection of the differences between ESCC miRNAs, the DGCR8$^{-/-}$ mESCs transfected with the control miRNA, were omitted in the PCA analysis (Figure 30b). In particular miR-291b and miR-294, clustered away from the other ESCC miRNAs.

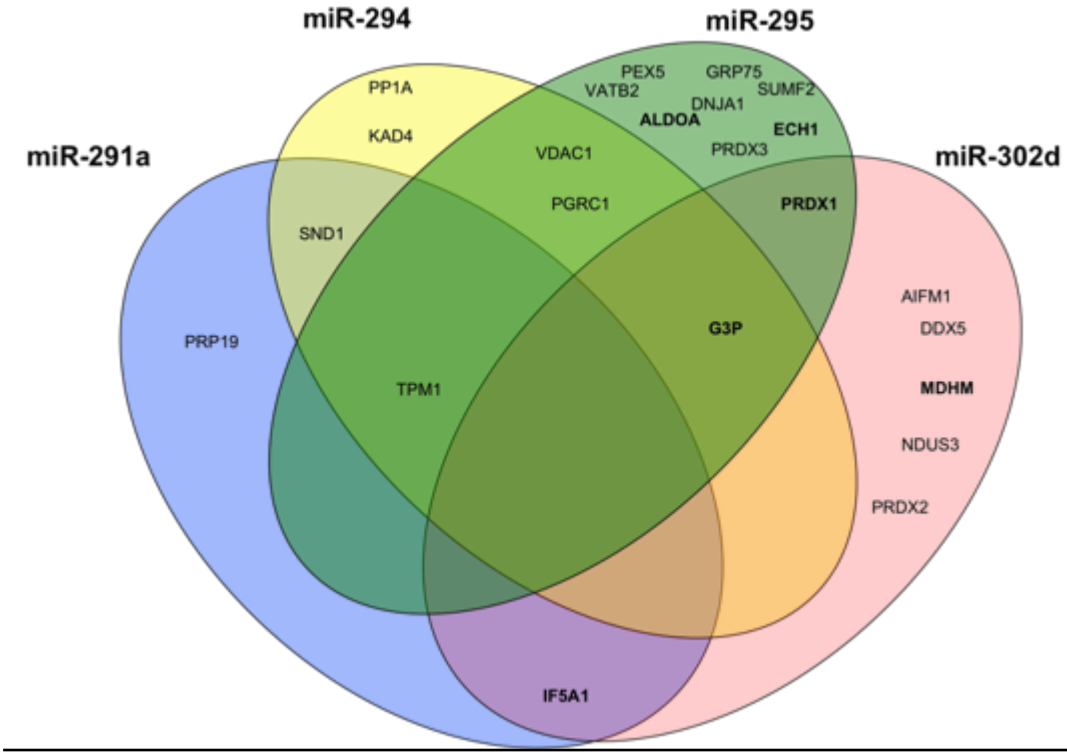


Figure 31 – Venn diagram of proteins from DIGE study. Proteins were selected based on significantly differential expression compared to cells transfected with the control miRNA. Proteins were selected based on having a fold expression difference of 1.2, with a significance level of p(ANOVA) <0.05. Proteins in bold represent proteins that were also identified as having differential expression between the wild type and DGCR8-/- mESCs at baseline using DIGE. Proteins are named using Uniprot nomenclature for mouse proteins (231).

Of the miRNAs studied (excluding miR-291b due its detrimental effect on cell number), miR-295 shared the most targets with the other miRNAs and also had the highest number of targets identified in both DIGE studies (Figure 31). miR-

291a, had the least number of targets of which three out of four of them were directly associated with transcription and translation (Table 14).

Twenty-three proteins were identified as being influenced by the expression of the ESCC miRNAs (Figure 31). Of these approximately a third were affected by the expression of at least two miRNAs, highlighting the high number of shared targets between miRNAs from the same family. Of these proteins, three were also identified in the initial DIGE study comparing baseline protein expression between the WT and DGCR8$^{-/-}$ mESCs; Peroxiredoxin 1 (PRDX1), Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and Eukaryotic translation initiation factor 5A-1 (IF5A1) (Figure 31, Table 14).

The identified proteins (Table 9) were searched in five different miRNA target predication databases, Targetscan (240), microRNA.org (241,242), DIANA (243,244), miRwalk (245) and miRNAmap (246) to uncover any predicted targets (Table 9).

Using the results from the DIGE analysis, we used Ingenuity pathway analysis software (237) to screen for any common pathways between proteins or any networks that the proteins may be involved in. The results are summarized in Figure 32.

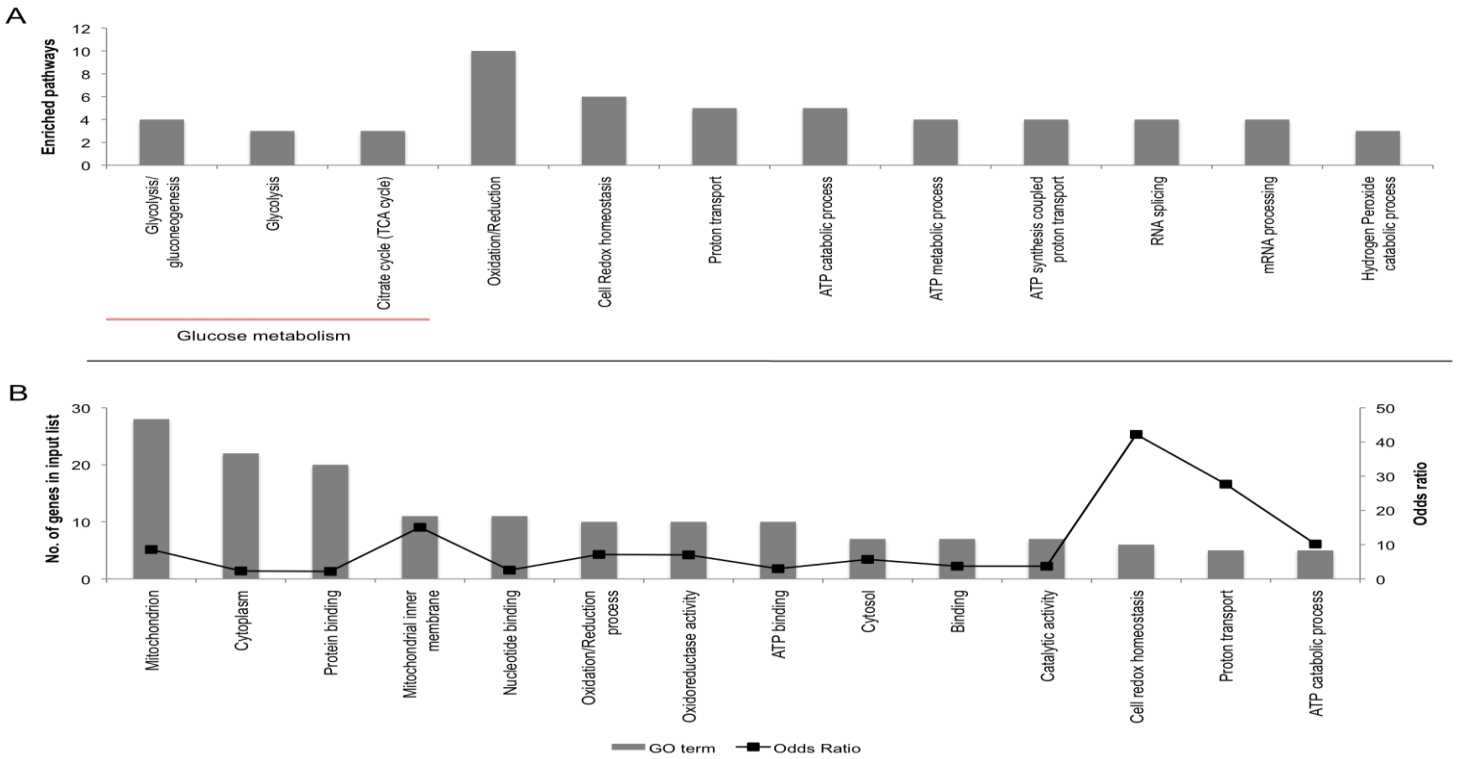


Figure 32- Pathway analysis. Only significant values are shown as determined by the DAVID software using a one-tailed Fisher's Exact test, with a cut-off value of $p<0.05$. A) Highly enriched pathways for proteins, which are differentially expressed upon individual transfections of the ESCC miRNAs. Only proteins with significant changes are represented as determined using a one-tailed Fisher's exact test, with a cut-off value of $p<0.05$. Image was created using the DAVID software (247), B) Proteins affected by the individual transfection of the ESCC miRNAs classified by GO terms B) Proteins affected by the individual transfection of the ESCC miRNAs, classified by GO terms. Height of bars is indicative of the number of significantly affected proteins classified by the GO term. Odds ratio=(no. of genes in input list/total list)/(no. of genes in the whole genome classified by the GO term/no. of genes in the whole genome).

The pathway analysis revealed that the most enriched pathways based on the transfection of the ESCCs, were those involved in redox reactions, proton transport and ATP metabolism (Figure 32a), which confirms the top GO term observed in Figure 32b, of 'Mitochondrion', and also the top odds ratio being aligned to the GO term, 'Cell redox homeostasis'.

In order to validate the results from the DIGE analysis (Figure 29, Table 9) as well as the accompanying common pathway analysis (Figure 32), western blots (Figures 33, 34) and qPCRs (Figure 35) were conducted on selected targets identified by proteomics, as well as associated mRNAs/proteins.

Initially, one way ANOVA tests were conducted on the mRNA and protein analysis of the transfected cells using GraphPad Prism (236) to determine if there were any significant differences in expression. If the p-value was less than 0.05, then further posthoc t-tests were conducted between the DGCR8$^{-/-}$ mESCs transfected with the individual ESCC and those transfected with the control miRNA to further assess any differences (Figures 32 -35).

9.2.3 Validation of proteins from DIGE analysis and associated proteins in transfected cells

Using ANOVA and individual t-tests to compare protein expression levels of DGCR8$^{-/-}$ mESCs transfected with the individual ESCCs and the mESCs transfected with a control miRNA, we were able to identify proteins influenced by the expression of the ESCC miRNAs (Table 9). Due to the similarities between the ESCC miRNAs particularly their shared seed sequence, many of the targets tested, were similarly affected (Figures 33- 35).

In order to validate the results from the DIGE analysis (Figure 29 and Table 14) as well as the accompanying common pathway analysis (Figure 32); Western blots (Figures 33 and 34) and qPCRs (Figure 35) were conducted not only analysing the proteins identified in the DIGE study but those in associated pathways, such as glycolysis and the TCA cycle.

Of the twelve mRNAs analysed due to their direct role in glycolysis, eight of them; displayed significant differential expression in the presence of the ESCC miRNAs compared to the DGCR8$^{-/-}$ mESCs transfected with the control miRNA (Figure 35a). Aldolase A in the presence of the ESCCs was further validated at the protein level (Figure 34a).

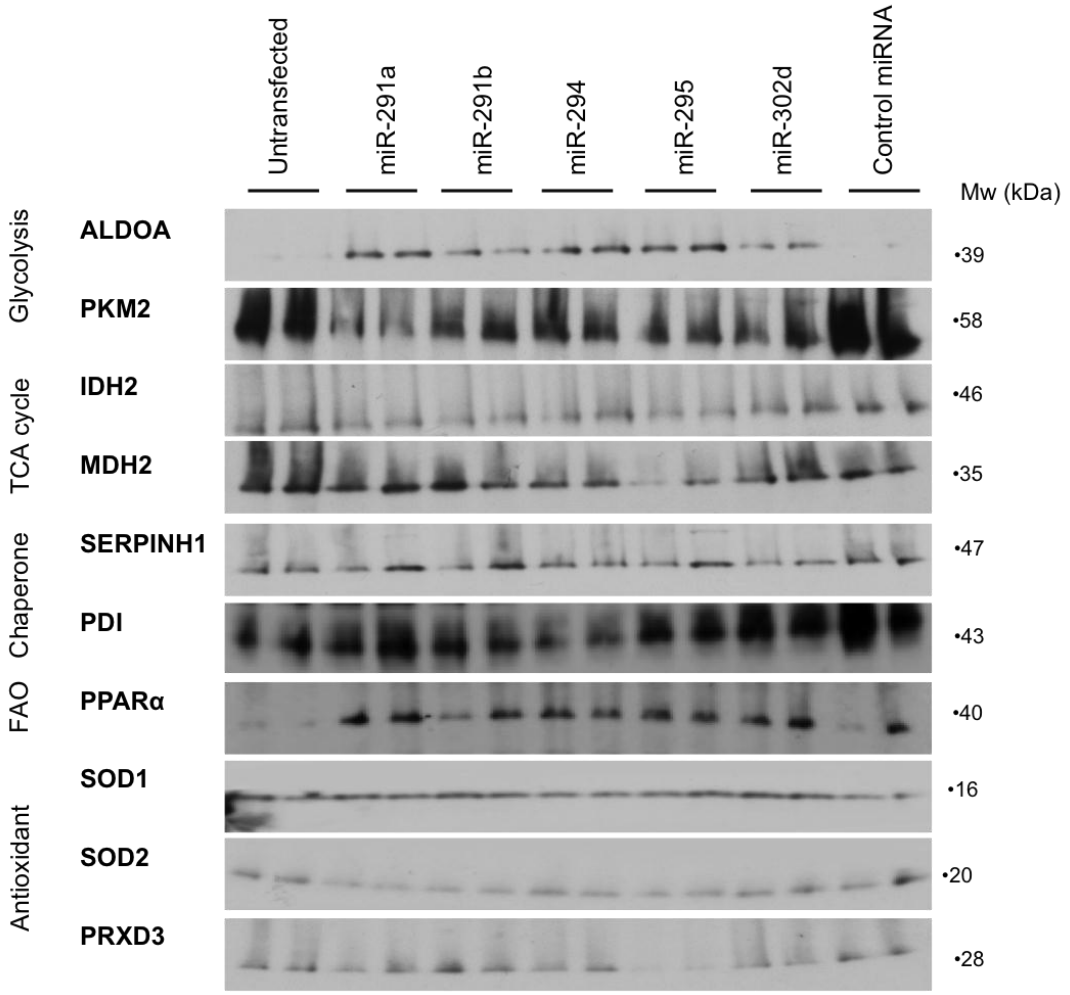


Figure 33 - Protein expression of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs. Immunoblotting was conducted on 3 consecutive passages of transfected cells. Analysed proteins were chosen based on their differential expression between the transfected cells (Table 9) and their relationship with identified proteins. Protein expression was normalized to Ponceau staining.

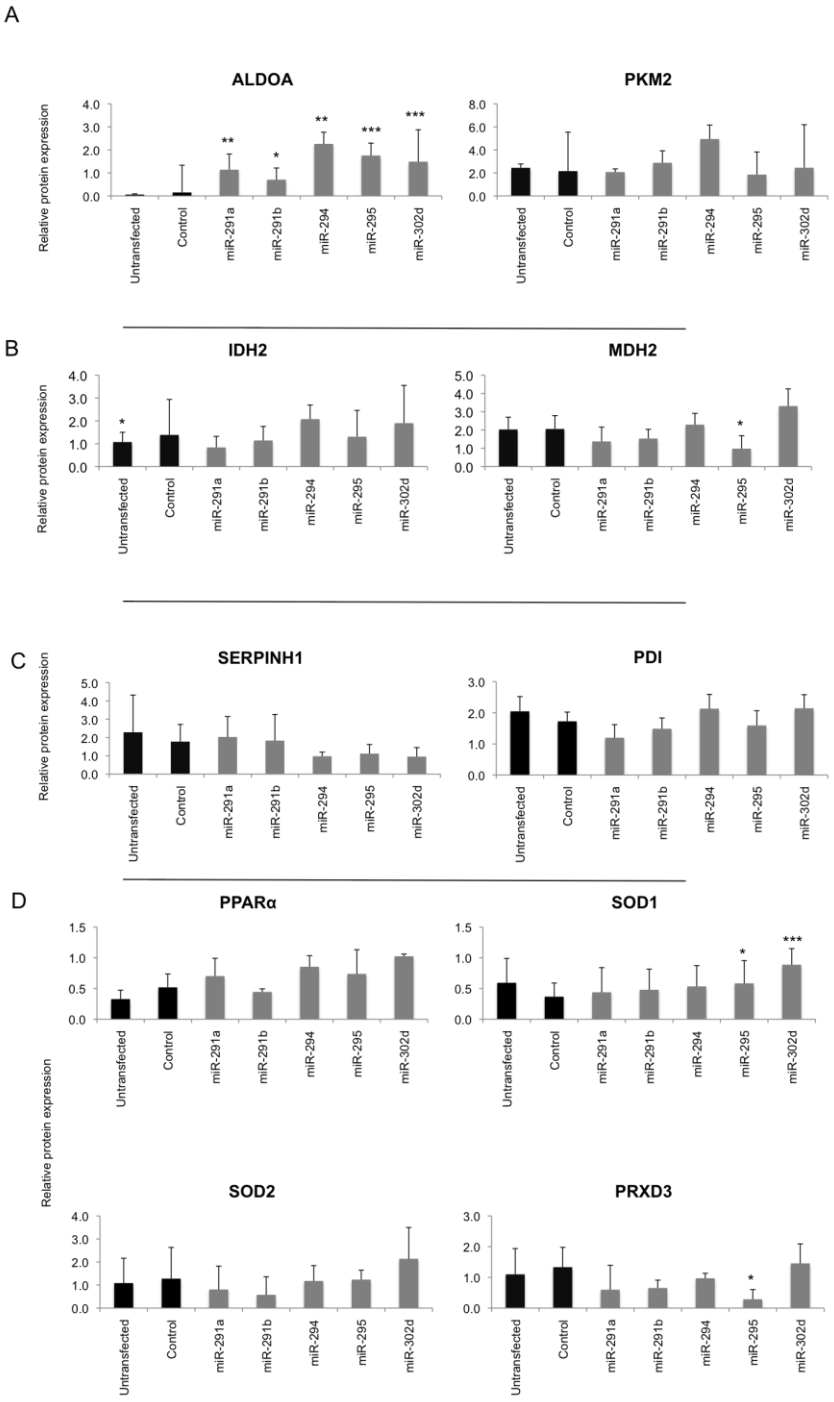
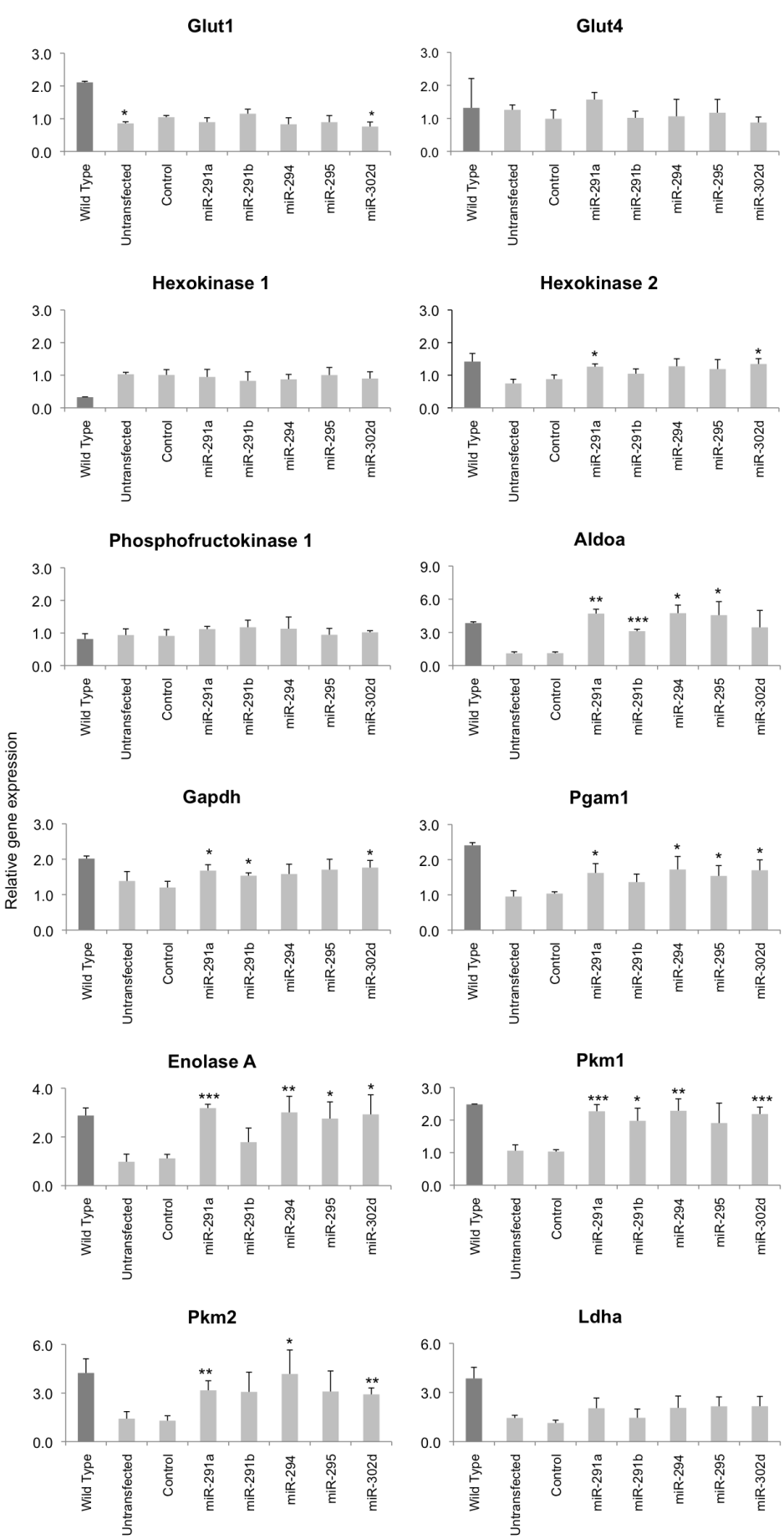
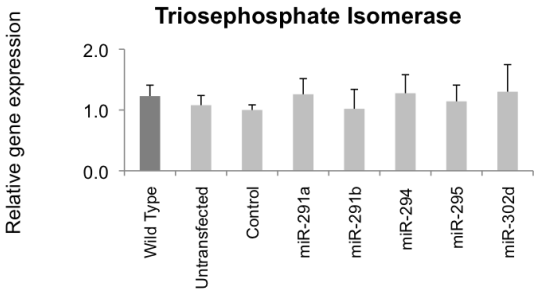


Figure 34 - Densitometry analysis of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs. A) Glycolytic proteins, B) TCA cycle enzymes, C) Chaperone proteins, D) Fatty acid metabolism and antioxidant protein. Protein expression levels were normalised to Ponceau staining. n=3, P(ANOVA)<0.05. *Abbreviations: ALDOA- Fructose bisphosphate aldolase A, PKM2- Pyruvate kinase M2 isoform, PDI- Protein disulfide isomerase, PPARα- Peroxisome proliferator-activated receptor α, SOD1/2- Superoxide dismutase 1/2, PRXD3- Peroxiredoxin 3*

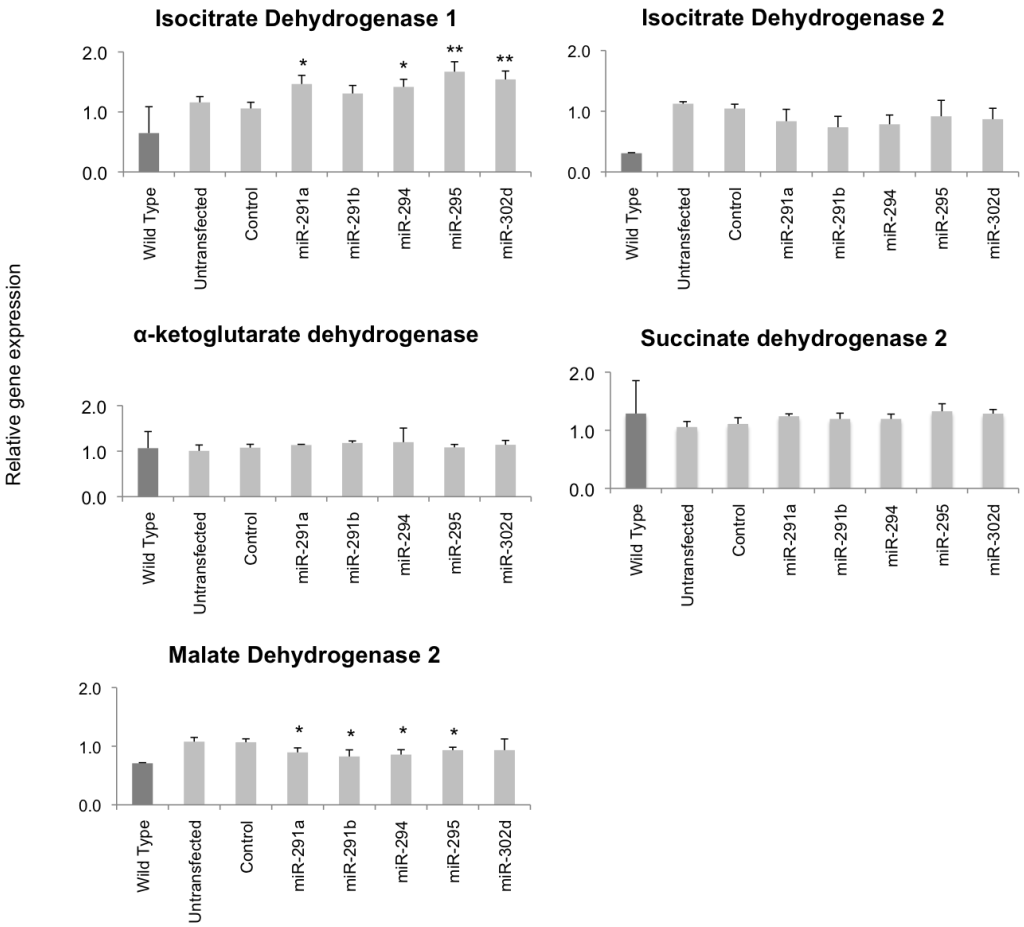
A


Glut1
Glut4
Hexokinase 1
Hexokinase 2
Phosphofructokinase 1
Aldoa
Gapdh
Pgam1
Enolase A
Pkm1
Pkm2
Ldha
Relative gene expression
Wild Type
Untransfected
Control
miR-291a
miR-291b
miR-294
miR-295
miR-302d

B


Triosephosphate Isomerase
Relative gene expression
2.0
1.0
0.0
Wild Type
Untransfected
Control
miR-291a
miR-291b
miR-294
miR-295
miR-302d

C


Isocitrate Dehydrogenase 1
Isocitrate Dehydrogenase 2
α-ketoglutarate dehydrogenase
Succinate dehydrogenase 2
Malate Dehydrogenase 2
Relative gene expression
2.0
1.0
0.0
*
**
Wild Type
Untransfected
Control
miR-291a
miR-291b
miR-294
miR-295
miR-302d

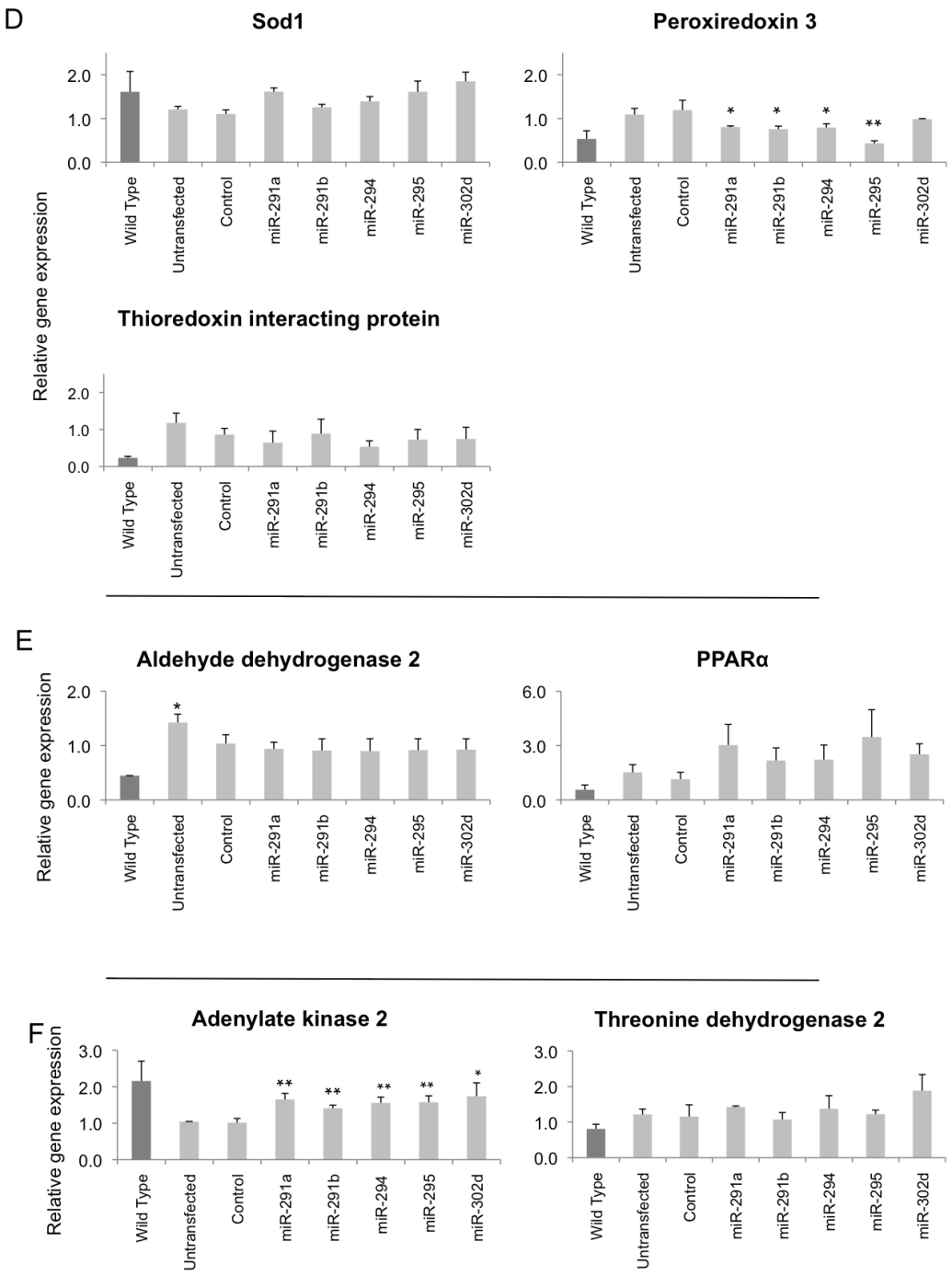


Figure 35 - Quantitative PCR profile of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs. A) Glycolytic genes, B) Gene associated with glycolysis, C) TCA cycle genes, D) Antioxidant genes, E) Fatty acid metabolism genes, F) Genes associated with energy metabolism. Gene expression levels were normalised to the loading control β-actin. n=3, P(ANOVA)=0.05. *Abbreviations- Aldoa- Fructose bisphosphate aldolase A, GAPDH-Glyceraldehyde-3-phosphate dehydrogenase, PGAM1-Phosphoglycerate mutase 1, PKM1/PKM2- Pyruvate kinase muscle 1/2*

The ability of the ESCC miRNAs to restore the expression of some of the glycolytic/TCA transcripts/proteins led us to further investigate the metabolomic status of the transfected cells, using ^{1}H-NMR techniques.

9.2.4 Metabolomics

In order to probe the effects of the individual ESCC on the metabolic status of mESCs, secretome analysis was conducted on the transfected DGCR8$^{-/-}$ cells (Figure 36).

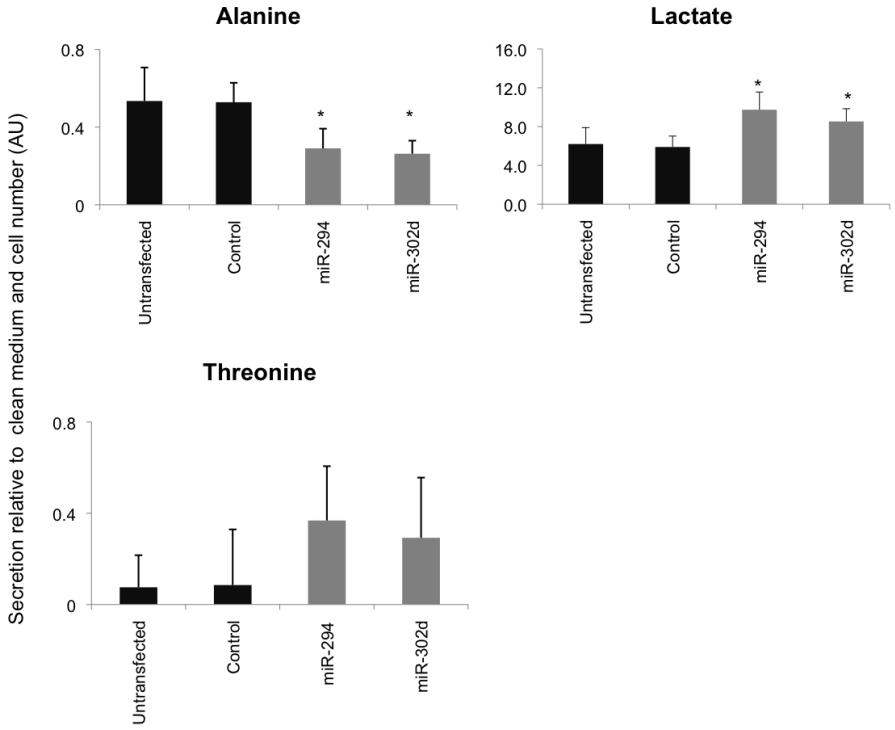


Figure 36– Secretome analysis of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs Metabolites were secreted into the media. Values were normalized to cell number n=3, P(ANOVA)<0.05

Although the DGCR8$^{-/-}$ mESCs were individually transfected with five miRNAs; miR-291a, miR-291b, miR-294, miR-295 and miR-302d, the secretome analysis revealed that metabolites showed the most significant differences upon the expression of miR-294 and miR-302d (Figure 36), which correlated with the principal component analysis of the proteomics data which displayed distinct separation from the other ESCC miRNAs (Figure 30). The analysis revealed numerous metabolites that are summarised in Table 12. Although miR-294 and miR-302d were the only significant miRNAs, all of the miRNAs tested affected alanine and lactate in a similar way, albeit not at a significant level. Therefore

further intracellular H^1 NMR analysis was conducted on the DGCR8$^{-/-}$ mESCs transfected with miR-294, miR-302d and the control miRNA (Figure 37).

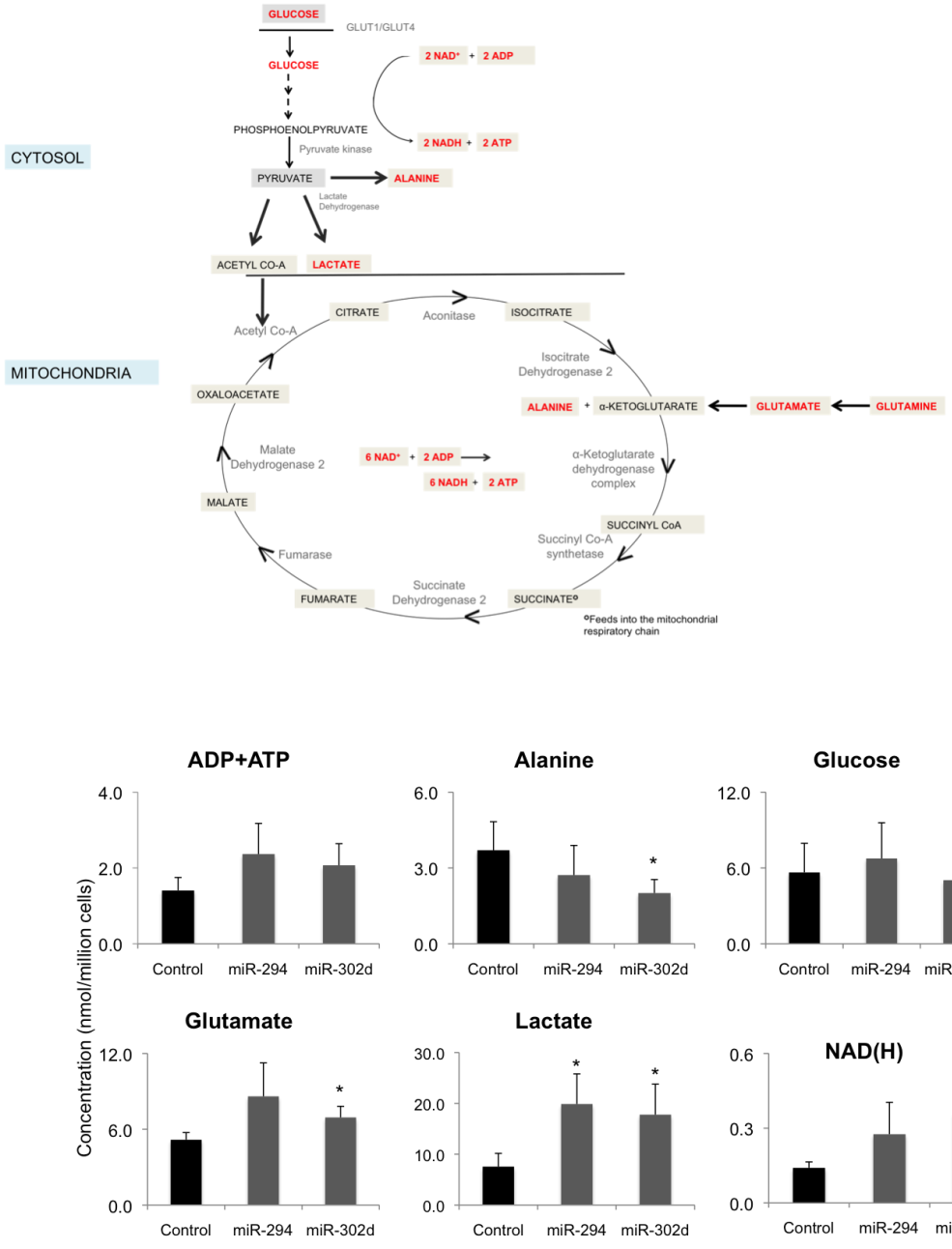


Figure 37 - Intracellular concentrations of metabolites in DGCR8$^{-/-}$ mESCs transfected with a control miRNA, miR-294 and miR-302d. Analysed metabolites highlighted in red. n=3, p(ANOVA)<0.05

The intracellular metabolite analysis showed that the DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs; miR-294 and 302d, had increased intracellular concentrations of lactate and decreased levels of alanine compared

to the DGCR8$^{-/-}$ mESCs transfected with the the control miRNAs (Figure 37). Notably, glutamate which can be metabolized in the mitochondria, into α-ketoglutarate (which can feed into the TCA cycle) displayed a higher intracellular concentration, in the transfected cells compared to the cells transfected with the control miRNA (Figure 37).

This is indicative of a possible switch from glycolysis to glutaminolysis for energy production (248) in the DGCR8$^{-/-}$ mESCs, which is shown to be partially reversed by the transfection of the ESCC miRNAs (Figures 32, 33).

In order to verify whether there is a switch from glucose metabolism to glutaminolysis in the DGCR8$^{-/-}$ cells for energy production, the gene expression of hypoxia inducible factor 1 α (HIF1α) (Figure 38) which can act as a master regulator for a switch from glycolysis to glutaminolysis (249,250) and glutamic pyruvate traminase 2 (*gpt2)*, the key enzyme which catalyses the reversible conversion of glutamate to alanine and α-ketoglutarate was analysed (Figure 39).

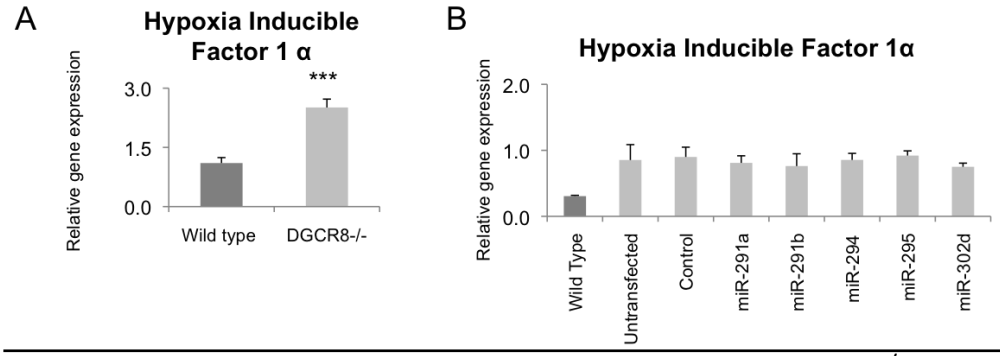


Figure 38 – mRNA expression of HIF1α. A) Expression in wild type and DGCR8$^{-/-}$ mESCs. Wild type expression levels were set to one, n=3, P(t-test)<0.05 B) Expression in DGCR8$^{-/-}$ mESCs transfected with ESCC miRNAs relative to DGCR8$^{-/-}$ mESCs transfected with a control miRNA. n=3, P(ANOVA<0.05)

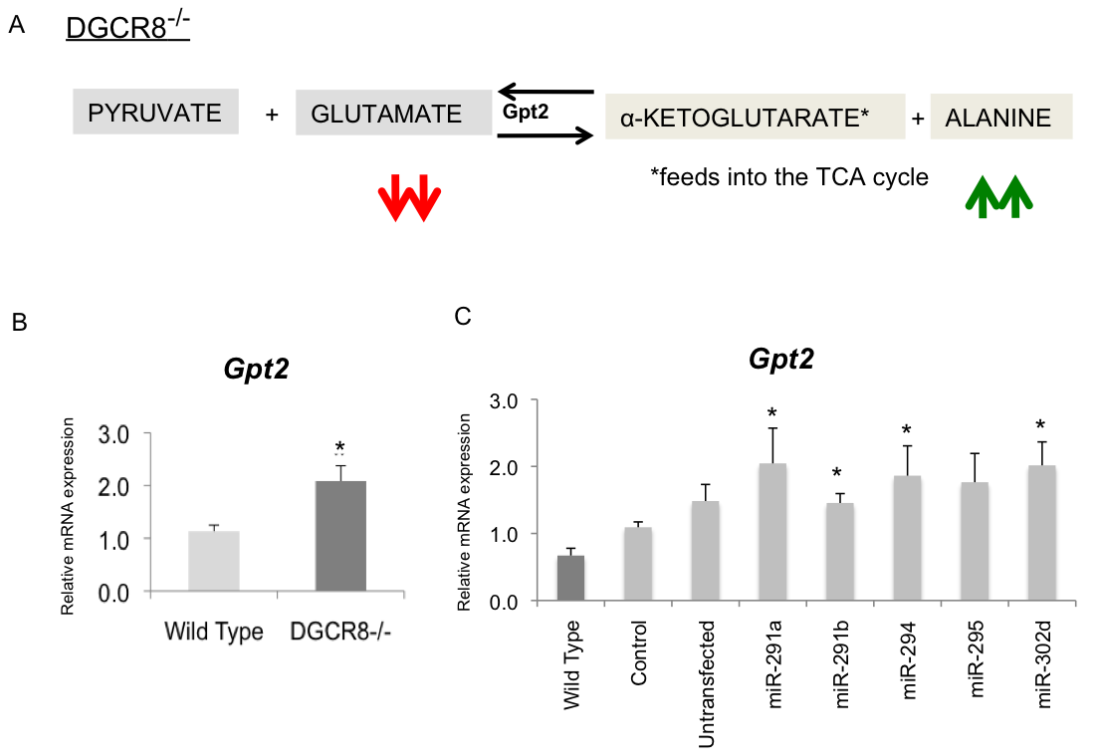


Figure 39 – Role and mRNA expression of glutamic pyruvate transaminase 2 *(gpt2)* in DGCR8$^{-/-}$ and WT mESCs. A) Summary of glutaminolysis pathway in DGCR8$^{-/-}$ mESCs (arrows represent intracellular metabolite levels relative to DGCR8$^{-/-}$ mESCs, B) Baseline gene expression levels of *gpt2* in both DGCR8$^{-/-}$ and WT mESCs, relative to the WT cells, n=3, P(t-test)=0.05, C) Quantitative PCR profile of *gpt2* in DGCR8$^{-/-}$ mESCs transfected with ESCC miRNAs relative to DGCR8$^{-/-}$ mESCs transfected with a control miRNA. n=3, P(ANOVA)<0.05

Both *Hif1α* and *Gpt2* showed significantly higher expression in the DGCR8$^{-/-}$ mESCs compared to the WT, and in the case of *Gpt2* this was enhanced by the overexpression of the ESCC miRNAs (Figure 34).

Taken together, the results of this study indicated that the absence of ESCC miRNAs in DGCR8$^{-/-}$ mESCs led to a metabolic switch from predominantly glucose to glutamine metabolism in these cells.

10. Discussion

In this study, we combined proteomics, metabolomics and RNA analysis techniques, to investigate the effect of the loss of DGCR8 and therefore miRNAs on mESCs. Using mass spectrometry techniques, we discovered a range of protein expression differences between the DGCR8$^{-/-}$ and WT mESCs, most notably there were consistent differences in the expression of glucose metabolism enzymes. Protein and RNA validation experiments, revealed there to be a distinct trend in the expression of glycolytic enzymes which displayed a lower expression in the DGCR8$^{-/-}$ mESCs compared to the WT with a reciprocal decrease in the expression of TCA cycle enzymes. This was supported by functional metabolic analysis, which revealed the WT mESCs to be more glycolytically active than the DGCR8$^{-/-}$ mESCs. By overexpressing highly conserved, mESC specific miRNAs, that are known to play an direct role in controlling the cell cycle and also play a role in the regulation of pluripotency, the expression of the majority of glycolytic enzymes was restored, indicating that the miRNAs tested influence the metabolic status of mESCs.

As miRNAs regulate gene expression at the post-transcriptional level, their effect is most keenly observed at the protein level, however studies on DGCR8$^{-/-}$ mESCs, predominantly focused on the transcriptome of these cells both at baseline and upon the transfection of specific miRNAs (8). No study to date has explored either the baseline proteome of DGCR8$^{-/-}$ mESCs, or the effects on the proteome of cells transfected with the ESCC miRNAs (miR-291a, miR-291b, miR-294, miR-295 and miR-302d). The over expression of these miRNAs in DGCR8$^{-/-}$ mESCs is sufficient to restore the cell cycle defects, yet it is unable to reverse the differentiation defects (3), however as there are clear links between the cell cycle and pluripotency, uncovering targets and networks affected by these miRNAs may provide insight into the regulation of the balance between pluripotency and differentiation (208).

It should be noted that proteomic techniques have been previously applied to the DGCR8 protein, in order to elucidate characteristics of its structure; tandem

mass spectrometry has been used to identify phosphorylation sites, which depending on the site in question, when phosphorylated can increase or decrease the stability of the DGCR8 protein, thereby acting as a regulatory mechanism for the successful formation of the microprocessor complex (176). Mass spectrometry has also been used to identify ribosomal proteins that bind to the 5' end of DGCR8 RNA, which may result in internal translation of DGCR8 in a cap-dependent manner (251).

10.1 Characterization studies

10.1.1 DGCR8$^{-/-}$ mESCs may display increased differentiation

Although both cell types displayed a similar genotype (Figure 14) ,after they had been cultured for 5 days, the cells displayed a distinctly different phenotype as indicated by not only their appearance but the different degrees of alkaline phosphatase staining (Figure 13). These results suggest that there is a profound difference in the cells underlying differentiation capacity, which is supported by Wang' seminal study on DGCR8$^{-/-}$ mESCs (3). This is supported by the initial DIGE analysis which indicated that as well as the DGCR8$^{-/-}$ mESCs expressing the pluripotency markers (Figure 14), but also expressing some early differentiation markers such as Calponin 3 and Vimentin, at a higher level than in the wild type cells (Table 7).

10.1.2 Sox2, but not Oct4 or Nanog display lower protein expression in the DGCR8$^{-/-}$ mESCs

In concordance with previous studies the WT and DGCR8$^{-/-}$ mESCs displayed similar levels of pluripotency, as defined by comparable levels of alkaline phosphatase staining and the high expression of the pluripotency markers; Oct4, Nanog and Sox2 in both cell types at the mRNA level (6). Yet only Oct4 and Nanog mirrored this trend at the protein level whereas, Sox2 displayed significantly lower expression in the DGCR8$^{-/-}$ mESCs compared to the WT. A possible reason for the discrepancy between the mRNA and protein expression of Sox2 may be the genetic structure of Sox2, which only consists of one exon.

Primers are usually designed to span at least two exons, in order to prevent DNA contamination from masking any results.

At the time of writing, there are no other studies, which compare pluripotency protein expression between DGCR8$^{-/-}$ and WT mESCs. DGCR8$^{-/-}$ mESCs were first characterized in the seminal study by *Wang et al., 2007,* which concluded that the three major pluripotency markers showed similar mRNA expression levels between the DGCR8$^{-/-}$ and WT mESCs under conditions that promote self renewal. Although the RNA samples were treated with DNase to prevent genomic DNA contamination, the primer set used to detect Sox2 was not specific and detected both Sox1 and Sox3, as well as Sox2 therefore any differences in Sox2 expression were not apparent (6).

Recently, a study comparing Sox2 protein expression between miR-295 cluster knockout and WT mESCs, has shown that the Sox2 protein is expressed at a similar level in both cell types, possibly signifying that miRNAs other than the ESCC miRNAs play a role in the regulation of Sox2 expression (252). This is supported by the protein analysis in this study that implies that Sox2 is not directly or indirectly regulated by the ESCC miRNAs, and studies, which show that although Oct4, Nanog and Sox2 bind to the promoter regions of the ESCC miRNAs sustaining their high expression, they are not targeted by the ESCC miRNAs (26). Notably, the miR-295 cluster has been implicated in the regulation of DNA methyltransferases (Dnmts), which upon differentiation aid in the permanent repression of Oct4 (253), therefore the ESCC miRNAs play an indirect role in maintenance of pluripotency by aiding in the onset of differentiation. This is supported by DGCR8$^{-/-}$ mESCs inability to downregulate markers of pluripotency even under differentiation inducing conditions (6).

The downregulation of Sox2 in the DGCR8$^{-/-}$ mESCs is at odds with the known phenotype of delayed or reduced differentiation rates in the DGCR8$^{-/-}$ mESCs (6,254). Particularly as Sox2 is known to maintain high levels of Oct4 expression, (49) and no difference is observed in the expression of Oct4 between both the WT and DGCR8$^{-/-}$ mESCs or upon the overexpression of the ESCC

miRNAs. As Sox2 shares redundant functions with other Sox proteins (44), and the downregulation of Sox2 in the DGCR8$^{-/-}$ mESCs does not have an effect on Oct4 levels as would be expected, it may indicate that the other isoforms of Sox2 may be playing a compensatory role or that miRNAs play a role in regulating the interaction between Oct4 and Sox2.

In WT mESCs, the expression of the pluripotency markers is tightly regulated to maintain the pluripotent state, this has been most reliably observed in studies where the expression of the pluripotency markers has been enhanced or repressed resulting in varying degrees of differentiation (37,45,49,50). In the particular case of Sox2, its downregulation results in differentiation towards the trophoectoderm lineage (53), however no difference was observed in spontaneous differentiation rates as confirmed by alkaline phosphatase staining or a decrease in Nanog or Oct4 expression levels, further highlighting the important role of the ESCC miRNAs in the maintenance of the balance between differentiation and pluripotency.

Interestingly, of the three major pluripotency markers, Sox2 is the only one, which is not specific to ESCs and is expressed by cells of the extra-embryonic ectoderm (52) and embryonic neural stem cells (53).

10.1.3 DGCR8$^{-/-}$ mESCs display a slower proliferation rate

mESCs have an unusual cell cycle structure that lack a G_1/S phase restriction point, caused in part by the high expression of the ESCC miRNAs (3). Due to the loss of miRNAs in DGCR8$^{-/-}$ mESCs, in particular the ESCC miRNAs, they have a significantly higher expression of cell cycle inhibitors of the G_1/S transition, including Rbl2, Lats and Cdkn1a (3). This results in a significantly slower proliferation rate compared to WT mESCs (3,6). We confirmed that the DGCR8$^{-/-}$ mESCs used in this study were displaying a slow growth phenotype and that the differences in cell number were not due to apoptosis which is in agreement with previous studies (6).

Recently studies have shown that the forced expression of the miR-295 and 302 clusters in DGCR8$^{-/-}$ mESCs, is enough to reverse the cell cycle/proliferation defect (3). In this study, the final cell numbers of WT and transfected DGCR8$^{-/-}$ mESCs was counted, and indicated that there was no significant difference between the two cell types, demonstrating that although the transfection efficiency levels were sufficient enough to affect known targets of the ESCC, it had no effect on the proliferation of the cells.

Initially conventional PCR was used to test the RNA expression levels of the cell cycle inhibitors, which revealed that Rbl2 and Lats2, displayed significantly higher expression in the DGCR8$^{-/-}$ mESCs compared to the WT. However, in stark contrast to the other tested cell cycle inhibitors and previous studies (3), although cdkn1a trended towards displaying lower expression in the DGCR8$^{-/-}$ mESCs, there was too much variability between the three technical replicates for there to be a significant difference between the two cell types. To eliminate any variability introduced by the insensitivity of the conventional PCR method, qPCR for the Cdkn1a mRNA was conducted on the same samples, and revealed that there was a significant difference in Cdkn1a expression between the two cell types. This was supported by the transfection studies, which confirmed that Cdkn1a is a direct target of the ESCC miRNAs.

Although all of the miRNAs had a similar effect in restoring the expression levels of cdkn1a, it should be noted that previous transfection studies have shown that of the ESCC miRNAs, miR-294 has the greatest effect on restoring the ESC cycle (3). This has led to it being termed a representative miRNA for the miR-295 cluster and it has been used as such, in transfection experiments studying the effects of the ESCC cluster (8,105,198,255)

10.1.4 The overexpression of miR-291b in DGCR8$^{-/-}$ mESCs results in increased cell death

The downregulation of DGCR8 results in increased susceptibility to cell death, limited proliferation and limited differentiation capacities, to varying degrees, in a range of cell types (256-259). Therefore changes in its expression levels have profound effects on cell phenotype. Nowhere, is this more distinct than in the DGCR8$^{-/-}$ mESC line, where the complete loss of DGCR8 results in reduced/delayed differentiation, slow proliferation and prolonged expression of the pluripotency genes, particularly; Oct4, Sox2 and Nanog under differentiating conditions (3,6). However studies have revealed that cell death rates do not differ between WT and DGCR8$^{-/-}$ mESCs (6), and this was confirmed in this study.

The miR-295 cluster has been shown to play an important role in protection against apoptosis; studies on Dicer knockout mESCs have shown that there is an upregulation of cell death genes such as *Caspase 2*, a direct target of the miR-295 cluster (148). However only when the Dicer knockout mESCs were exposed to DNA damaging chemicals, did the protective properties of the miR-295 cluster become apparent (148). Interestingly, DGCR8$^{-/-}$ mESCs do not downregulate pro-survival, C-terminal binding proteins due to the loss of miRNAs, when treated with an apoptotic reagent such as Staurosporine, therefore the loss of miRNAs can protect against apoptosis (260), thus miRNAs are not only important for controlling mESC proliferation via the cell cycle but also for regulating mESC apoptosis (3,208).

Although, there were no differences in cell death between the WT and DGCR8$^{-/-}$ mESCs at baseline, and on average post-transfection of the ESCC miRNAs there were no differences in cell number, in two of the four replicates used to study cell proliferation in the cells transfected with miR-291b, there was a 5% loss of cells compared to other transfections (Table 11). This may have explained the phenotypic differences between the cells transfected with miR-291b, which displayed less round colony formation and more elongated cells around the edges of colonies. This correlates with a previous study, which shows that the

remaining miRNAs in the 295-cluster with the specific seed sequence 'AAGUGC', have a protective effect with regards to cell death when re-introduced into miRNA deficient mESCs (148). Another contributing factor may have been the considerably higher transfection efficiency of miR-291b in the DGCR8$^{-/-}$ mESCs compared to the other transfections, as well as the one nucleotide difference in seed sequence in miR-291b compared to the other tested miRNAs possibly indicating that as well as miR-291b targeting the same mRNAs as the other ESCCs, it also has other distinct mRNA targets, which result in the morphological and cell number differences.

The characterization experiments conducted on the WT and DGCR8$^{-/-}$ mESCs indicate that the loss of DGCR8 results in lower protein expression of Sox2 and a slower proliferation rate in comparison to WT cells, which has been attributed at least in part to the loss of miRNAs from the miR-295 cluster, resulting in decreased expression of the cell cycle regulators (3). These results concur with the seminal study by *Wang et al*., which show that the differences in proliferation observed between DGCR8$^{-/-}$ and WT mESCs is solely due to a decreased proliferation rate, at least in part caused by an accumulation of cells in the G_1/S phase of the cell cycle. (6),

10.2 Proteins influenced by the ESCC miRNAs

We chose to study the targets and effects of the miRNAs from the miR-295 cluster; miR-291a, miR-291b, miR-294 and miR-295 and a human homolog miR-302d (ESCC miRNAs), as the miR-295 cluster is the most abundantly expressed group of miRNAs in mESCs (48), is known to control the ESC cycle and has also been instrumental in increasing the efficiency of iPSC formation, in comparison to the Yamanaka factors (16,214), and therefore play a major role in the maintenance and regulation of pluripotency (3). Due to the importance of the ESCC miRNAs for the stem cell phenotype and their high abundance in mESCs, it is imperative that the pathways that these miRNAs influence are studied.

Of the miRNAs studied (excluding miR-291b, due to it occasionally promoting cell death in our study), miR-295 shared the most targets with the other miRNAs and also had the highest number of targets identified in both DIGE studies. miR-291a, had the least number of targets of which three out of four were directly associated with transcription and translation indicative of miR-291a playing a general role in regulating gene expression.

Fifty-eight proteins showed differential expression between the different transfections and of these twenty-three proteins were identified as showing significantly differential expression compared to the mESCs transfected with the control miRNA. Of these approximately a third were affected by the expression of at least two miRNAs, highlighting the high number of shared targets between miRNAs from the same family. Furthermore three of these proteins; Aldolase A, Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and Peroxiredoxin 1 were also identified in the initial DIGE study comparing baseline protein expression between the WT and DGCR8$^{-/-}$ mESCs. Interestingly, both Aldolase A and GAPDH are key enzymes for glycolysis, and upon overexpression of the ESCC miRNAs, the expression of both of the enzymes was restored at the mRNA.

Interestingly, there is emerging evidence alluding to the fact that metabolism may play a role in the control of differentiation/maintenance of pluripotency (24,60,87,261,262). Although the proposed links between glucose metabolism and ESC pluripotency have not been directly identified, studies have shown that pluripotent cells predominantly use glycolysis as their primary form of energy generation, and upon differentiation switch to predominantly using oxidative phosphorylation (28,62). Also, during the reprogramming process from mature cell types to iPSCs, glycolytic genes are switched on before pluripotent genes, indicating that the switch in glucose metabolism may be a trigger for the onset of pluripotency (57,63,70). However these studies do not directly and definitively uncover the mechanisms that connect pluripotency and metabolism, the results of the DIGE analysis potentially indicate that it is miRNAs that control glucose metabolism in the mESCs, and this may influence the pluripotent state.

In the following chapter, I will discuss glucose metabolism in DGCR8$^{-/-}$ mESCs, focusing on glycolysis (Aldolase A, Phosphofructokinase 1, PKM2 and Fructose-1,6-bisphosphate) and the TCA cycle (Succinate dehydrogenase 2, Isocitrate dehydrogenase 2 and Malate dehydrogenase 2).

10.3 Glucose metabolism

10.3.1 DGCR8$^{-/-}$ and WT mESCs display differential expression of glucose metabolism enzymes

The DIGE analysis revealed that one of largest groups of proteins that showed differential expression between the WT and DGCR8$^{-/-}$ mESCs, were those involved in glucose metabolism. This was supported by the principal component analysis, which indicated that the top three pathways containing the most differentially expressed proteins were involved in glucose metabolism related pathways; 'Gluconeogenesis', 'Glucose metabolism' and 'Glycolysis'. Another metabolic pathway that was identified was the 'Citric acid cycle (TCA cycle)'. The differential expression of enzymes involved in both glycolysis and the TCA cycle was validated at both the RNA and protein levels. Strikingly, the differences observed in the expression of the glucose metabolism genes were more pronounced at the protein level compared to the RNA level suggesting that glucose metabolism genes may also be regulated at the post translational level. This was also supported by the common pathway analysis conducted on the DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs, which revealed that the top three enriched pathways were those involved in redox reactions, proton transport and ATP metabolism, which is tightly interlinked to glucose metabolism (Figure 32).

As well as the top three pathways, the pathway analysis also identified 'Translocation of GLUT4 to the plasma membrane' and 'Translocation of GLUT4 vesical and docking at the plasma membrane' as common pathways for the differentially expressed proteins between the WT and DGCR8$^{-/-}$ mESCs, which

corresponds to the mRNA expression pattern of the glucose transporters GLUT1 and even more importantly, GLUT4. Among the well-characterized family of glucose transporters, only GLUT1 and GLUT4 were analysed as they specifically shuttle glucose into cells (263) and both displayed significantly lower mRNA expression in the DGCR8$^{-/-}$ compared to the WT mESCs.

The pathway analysis was mirrored with eight of the twelve glycolytic enzymes analysed, which all showed significantly lower mRNA expression in the DGCR8$^{-/-}$ compared to the WT mESCs, suggesting that they are indirect targets of the ESCC miRNAs. Of the four glycolytic mRNAs that did not show significantly different expression between the two cell types, Phosphofructokinase 1 (pfk1) was the most notable, as it fully commits cells to glycolysis by catalysing the second irreversible step (Berg, Tymoczko et al. 2002), and the remaining rate-limiting enzymes; hexokinase 2 and pyruvate kinase 1/2 were expressed at lower levels in the DGCR8$^{-/-}$ mESCs. This indicates that pfk1 is not affected by the ESCC miRNAs and this is confirmed in the transfection experiments, possibly alluding to the fact that pfk1's activity is regulated by a number of binding factors including a high ATP/AMP ratio, a low pH as can be caused by high lactate production and high citrate production which act as inhibitors and fructose-2, 6-bisphosphate which acts as an activator (68). The binding of these proteins results in a conformational change, affecting the activity of the enzyme, however the mRNA analysis only provides data on gene expression not enzymatic activity.

PKM2 which also exists in two states, as both a low activity dimer and high activity tetramer form (264), which could not be distinguished between using PCR or immunoblotting, displayed lower expression in the DGCR8$^{-/-}$ mESCs. Interestingly, although the M2 isoform is predominant form of the PKM2 enzyme in mESCs (74), the M1 form also showed significantly differential expression between the two cell types, supporting a higher rate of glycolysis in the WT cells compared to the DGCR8$^{-/-}$.

This was further supported by Seahorse analysis, which provided a functional read out of the glycolytic rates of the DGCR8$^{-/-}$ and WT mESCs, and revealed that

although at baseline both cell types had a similar ECAR (indicative of the rate of glycolysis), upon the introduction of Oligomycin, an ATP synthase inhibitor, which forces the cells to maximally use glycolysis for energy generation; the WT cells displayed a higher glycolytic capacity. This coupled with the RNA and protein analysis, possibly indicates that although at baseline both cells conduct glycolysis at a similar rate, when mitochondrial function is impaired, the WT cells are poised (via the higher expression of glycolytic enzymes) to use glycolysis more effectively for energy generation.

Lactate, which is an end product of glycolysis, was secreted in higher quantities upon the overexpression of the ESCC miRNAs in comparison to the DGCR8$^{-/-}$ mESCs transfected with the control miRNA, which correlated with the higher expression of LDHA in the WT cells compared to the DGCR8$^{-/-}$. These results suggest that the ESCC miRNAs play a role in positively controlling the rate of glycolysis in wild type mESCs.

To verify this a wide range of glycolytic enzymes were analysed upon the overexpression of the ESCC miRNAs in the DGCR8$^{-/-}$ mESCs. Seven of the twelve mRNAs analysed, expression was restored near to or to WT levels. Notably both Aldolase A and PKM2 displayed restoration of mRNA expression to WT levels, and strikingly this was mirrored at the protein level for Aldolase A. Interestingly studies have shown that nuclear PKM2 acts in a positive feedback loop to promote the higher expression of glycolytic enzymes, such as GLUT1 and lactate dehydrogenase (86,265).

Although not all of the miRNAs were able to restore expression to levels that were significant compared to the mESCs transfected with the control miRNA, they all followed a similar trend, confirming that due to the shared seed sequences of the miRNAs, they share a large number of targets.

The metabolic analyses accompanied by the lower expression of glycolytic proteins and mRNA in the DGCR8$^{-/-}$ cells, indicate that the DGCR8$^{-/-}$ cells are less

glycolytic than the WT cells, and that this is influenced by the expression of the ESCC miRNAs.

10.3.2 Aldolase A's substrate FBP regulates PKM2 expression

PKM2, the rate-limiting enzyme for the final step of glycolysis, exists in either a less active dimeric form or a highly active tetrameric form that promote anabolic processes which debranch from glycolysis or glycolysis, respectively (80). Although the qPCR and protein analysis do not indicate what form of the PKM2 enzyme is most dominant in either of the cell types, they do indicate that overall PKM2 is expressed at lower levels in the DGCR8$^{-/-}$ compared to the WT cells, and that PKM2 is an indirect target of miR-291b, miR-294, miR-295 and miR-302d, at the mRNA level. However, the upregulation of PKM2 in the presence of the ESCCs at the RNA level, was not mirrored at the protein level for reasons that are yet to be elucidated.

The intraconversion of PKM2 between its dimeric and tetrameric forms is controlled by the binding of activators such as fructose-1, 6-bisphosphate (FBP) and repressors, including tyrosine-phosphorylated peptides. Notably, FBP is a major substrate for Aldolase A during glycolysis, therefore the expression of higher expression of Aldolase A in the WT mESCs compared to the DGCR8$^{-/-}$ coupled to a higher amount of lactate production in the presence of the ESCC miRNAs in the transfected DGCR8$^{-/-}$ mESCs, further supports a higher rate of glycolysis in the WT cells compared to the DGCR8$^{-/-}$ (Figure 40).

The differential expression of Aldolase A between the DGCR8$^{-/-}$ and WT mESCs, as well as the restoration of Aldolase A to WT levels upon the introduction of the ESCC miRNAs into the DGCR8$^{-/-}$ mESCs suggests that the expression of Aldolase A may be indirectly regulated by the ESCC miRNAs. Notably, high levels of Aldolase A expression in cells has been correlated to a high rate of glycolysis in mESCs (266).

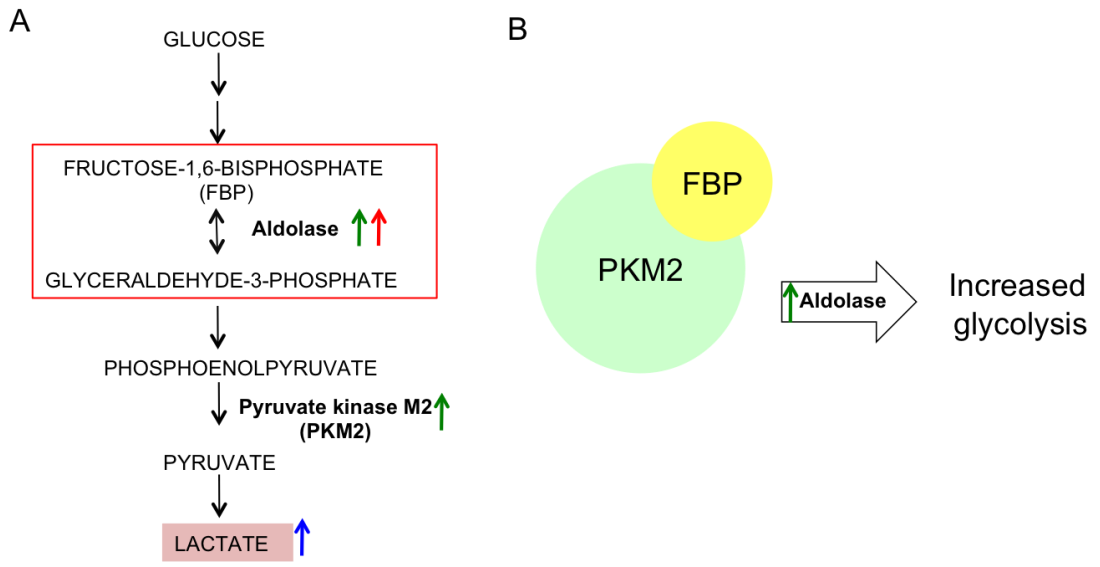


Figure 40- Simplified representation of glycolytic pathway in WT mESCs A) Green arrows represent mRNA expression compared to DGCR8$^{-/-}$ mESCs. Red arrow represents protein expression compared to DGCR8$^{-/-}$ mESCs. Blue arrow represents lactate secretion compared to DGCR8$^{-/-}$ mESCs. B) Binding of FBP to PKM2 results in higher glycolysis via higher expression of Aldolase A.

10.3.3 The DGCR8$^{-/-}$ mESCs display a higher expression of TCA cycle enzymes

Reciprocal to the expression pattern of the glycolytic genes, the DIGE analysis revealed that the TCA cycle proteins showed a trend of being more highly expressed in the DGCR8$^{-/-}$ mESCs compared to the WT, with only one out of the six proteins identified being expressed at higher levels than in the DGCR8$^{-/-}$ mESCs. This was reflected in the common pathway analysis, which indicates that within DGCR8$^{-/-}$ mESCs, the individual ESCCs may play a role in the regulation of proteins involved in the TCA cycle and also in the regulation of oxidative stress. Succinate dehydrogenase assembly factor 2 which aids in the assembly of the succinate dehydrogenase 2 enzyme complex (267) was the one TCA cycle associated protein that was identified as showing higher expression in the WT mESCs compared to the DGCR8$^{-/-}$, this was mirrored by RNA analysis for succinate dehydrogenase 2. Interestingly, succinate dehydrogenase 2 couples the TCA cycle to oxidative phosphorylation, possibly indicating that although there is an upregulation of TCA cycle proteins, the coupling between oxidative phosphorylation and the TCA cycle is weaker in the DGCR8$^{-/-}$ mESCs, suggesting

that the DGCR8$^{-/-}$ mESCs may be using another metabolic pathway for energy generation.

The DIGE analysis revealed that Isocitrate Dehydrogenase 2 (Idh2); which catalyses the sole rate limiting step of the TCA cycle (268) showed one of the largest fold changes of 2.25 higher in the DGCR8$^{-/-}$ than the WT mESCs and this was confirmed at both the RNA and protein levels. However, in this study Idh2 was not identified as a target of any of the miRNAs, though the qPCR analysis revealed it to be downregulated upon transfection of miR-291b, with a p-value on the cusp of significance at 0.051.

Malate dehydrogenase 2 (Mdh2), another TCA cycle enzyme which also showed significantly higher expression in the DGCR8$^{-/-}$ in comparison to the WT mESCs at the protein level, was directly affected by the expression of the ESCC miRNAs, further alluding to a role of the ESCC miRNAs in controlling the metabolic phenotype of mESCs. Although all of the miRNAs downregulated the expression of Mdh2 at the RNA level, miR-295 had the most obvious effect on Mdh2 expression at the protein level, possibly indicating that although Mdh2 is expression is downregulated by all of the ESCC miRNAs, miR-295 has the greatest effect on its expression. This is supported by miRNA target prediction databases that show Mdh2 to be targeted by miR-291a and 291b (245).

Seahorse analysis was used to provide functional data to support the mRNA and protein expression data. At baseline, the DGCR8$^{-/-}$ cells had a significantly lower OCR indicative of a lower rate of mitochondrial respiration at baseline, which may be representative of their slower proliferation rate compared to the WT mESCs. This was also the case upon the addition of Oligomycin A, an ATP synthase inhibitor, which indicated a lower amount of ATP production by the DGCR8$^{-/-}$ mESCs.

The maximal respiration rate was calculated using the OCR values upon the addition of FCCP minus the non mitochondrial respiration (values collected upon the addition of Rotenone and Antimycin A), however although the WT showed a

higher maximal respiration rate overall, the difference was not significant indicating that both cell types have the capacity for similar mitochondrial function.

Overall, these results indicate that although both the WT and DGCR8$^{-/-}$ mESCs are still reliant on glycolysis, as indicated by a similar basal glycolytic rate, the lower expression of the glycolytic enzymes accompanied by the higher expression of some TCA cycle proteins indicates that the DGCR8$^{-/-}$ mESCs may also be using an alternative form of energy generation.

10.4 DGCR8$^{-/-}$ mESCs may use an alternative energy generation pathway to glycolysis

This study suggests that DGCR8$^{-/-}$ mESCs are less glycolytic than their WT counterparts. Metabolic techniques were used to discover whether the DGCR8$^{-/-}$ mESCs were using an alternative pathway for energy generation. The analysis showed that the DGCR8$^{-/-}$ mESCs displayed a lower expression of glycolytic enzymes, and the expression of the majority of glycolytic enzymes studied was restored upon the expression of the ESCC miRNAs in the DGCR8$^{-/-}$ mESCs. This was supported by both NMR based secretome and intracellular analysis, which revealed the DGCR8$^{-/-}$ mESCs to contain and secrete less lactate, and also secrete less alanine, which was reversed by the expression of the ESCC miRNAs. Threonine secretion was also increased in the presence of the ESCC miRNAs. The analysis suggested that similar to cancer cells, the DGCR8$^{-/-}$ mESCs may be using glutaminolysis as a compensatory energy source and that changes in threonine metabolism may be playing in role in the survival of DGCR8$^{-/-}$ mESCs.

Glutaminolysis is defined as '...a series of biochemical reactions by which glutamine is lysed to glutamate, aspartate, CO_2, pyruvate, lactate, alanine and citrate...' (269), thereby producing intermediates such as glutamate and aspartate which can be used as precursors for nucleic acid and serine synthesis, respectively (270). It is a key energy source for proliferative cells, particularly

cancer cells, as it allows them to produce ATP under low glucose conditions (271).

In comparison to the WT mESCs, the DGCR8$^{-/-}$ cells displayed a lower expression of glycolytic enzymes and a reciprocal higher expression of TCA cycle enzymes. The metabolic analyses, show that the transfected DGCR8$^{-/-}$ mESCs have a lower intracellular concentration of alanine and a higher intracellular concentration and lactate compared to the DGCR8$^{-/-}$ mESCs transfected with the control miRNAs and reciprocally secrete less alanine and more lactate. Notably, glutamate and alanine displayed a lower and higher intracellular concentration, respectively, in the transfected cells compared to the cells transfected with the control miRNA. Glutamate can be metabolized in the mitochondria, into α-ketoglutarate which can feed into the TCA cycle. This is indicative of a switch from glycolysis to glutaminolysis for energy production (248) in the DGCR8$^{-/-}$ mESCs, which is shown to be partially reversed by the transfection of the ESCC miRNAs.

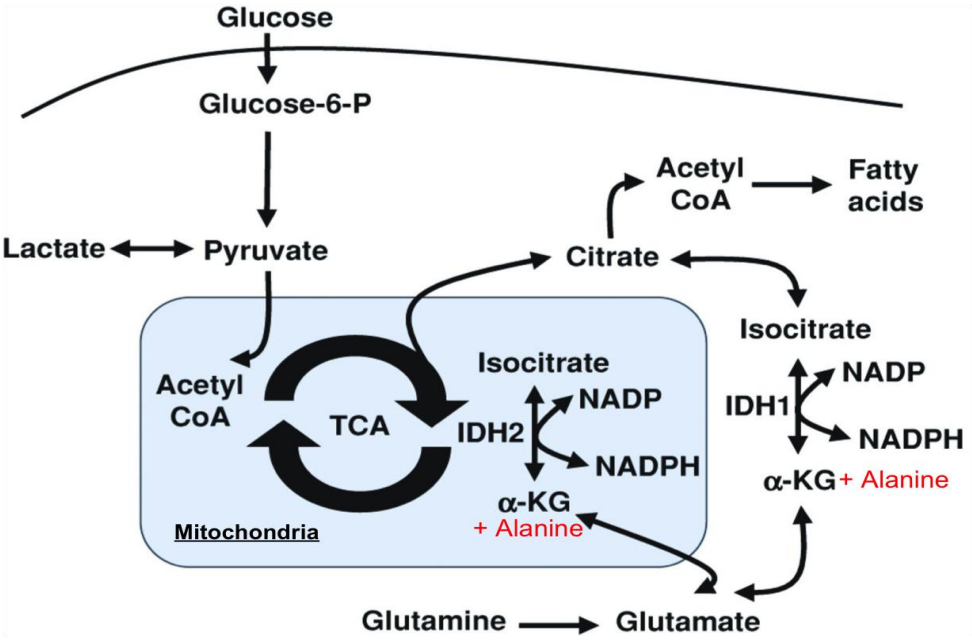


Figure 41- Glutaminolysis via the 'reverse TCA cycle' in rare cancer cells. Glucose is converted into pyruvate, which is converted into acetyl Co-A that feeds into the reverse TCA cycle. Glutamine is metabolized into α-ketoglutarate (α-KG) and alanine, and then Isocitrate via the high expression of the Isocitrate dehydrogenase isozymes (IDH), resulting in a higher production of fatty acids. Adapted from: (272)

Interestingly, studies have shown that cancer cells (Figure 41) which are similar to mESCs in terms of the expression of pluripotent genes and their ability to use aerobic glycolysis for energy generation, can in specific cases undergo metabolic changes which result in them using glutaminolysis via 'the reverse TCA cycle' aided by the ability of the Isocitrate dehydrogenase enzymes to catalyse reversible reactions, for energy generation, thereby producing less lactate and also resulting in higher fatty acid production (273).

This is supported by the higher baseline expression of TCA cycle proteins, most significantly Idh2 in the DGCR8$^{-/-}$ mESCs compared to the WT, which correlates to the higher expression of the Idh isoforms in certain cancer types (271,273), the increased protein expression of fatty acid related proteins as identified by the DIGE analysis, as well as the higher expression of HIF1α in the DGCR8$^{-/-}$ mESCs, which under hypoxic conditions has been found to be a master regulator of the switch to glutaminolysis in some rare cancer types (249).

To aid with the verification of this, the expression of Pparα, a transcription factor/ master regulator, which upon activation promotes uptake, utilization and catabolism of fatty acids (274) was analysed at the protein and RNA levels. The blot shows that there is a consistent trend of Pparα being more highly expressed in the DGCR8$^{-/-}$ mESCs compared to the wild type, with the p-value being close to significance, at 0.06, at the RNA level the difference is significant (p-value= 0.021). Aldehyde dehydrogenase 2, another protein involved in fatty acid metabolism, of which two isozymes were identified in the DIGE analysis as being differentially expressed between the wild type and DGCR8$^{-/-}$ mESCs is a predicted target of the ESCC miRNAs as identified in three miRNA data prediction databases (240,243,244) and was more highly expressed in the DGCR8$^{-/-}$ compared to the WT cells. However both Pparα and Aldehyde dehydrogenase 2, were unaffected by the presence of the ESCC miRNAs.

Triosephosphate Isomerase (TPI) which is also involved in glycolysis and is known to reversibly catalyse the conversion of Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) to Dihydroxyacetone phosphate (DHAP) exhibited

similar expression in both the WT and DGCR8$^{-/-}$, notably DHAP is known to be a precursor for fatty acid synthesis. However the expression of TPI was similar across both cell types, possibly indicating that if fatty acid metabolism rates have increased in the DGCR8$^{-/-}$ mESCs, then this is due to an alternative pathway such as the reverse TCA cycle summarized in Figure 41.

Most importantly the NMR analysis revealed that at baseline, alanine is excreted at higher concentrations in the DGCR8$^{-/-}$ mESCs compared to the WT and upon the overexpression of ESCC miRNAs, the secretion levels decreased indicating the miRNAs have a direct effect on alanine metabolism. Lactate displayed higher secretion levels upon overexpression of ESCC miRNAs in the DGCR8$^{-/-}$ mESCs.

Although the DGCR8$^{-/-}$ mESCs were individually transfected with five miRNAs; miR-291a, miR-291b, miR-294, miR-295 and miR-302d, the secretome analysis revealed that the metabolites displayed the most significant differences upon overexpression of miR-294 and miR-302d compared to cells transfected with the control miRNA, which correlated with the principal component analysis. Therefore intracellular H^1 NMR analysis was conducted on the DGCR8$^{-/-}$ mESCs transfected with miR-294, miR-302d and the control miRNA.

10.4.1 HIF1α expression in DGCR8$^{-/-}$ mESCs

Notably although HIF1α is a predicted target of the ESCC miRNAs (except miR-291b, possibly due to the one nucleotide difference in seed sequence), it was unaffected by the presence of the ESCC miRNAs, indicating that the activation of PKM2 via HIF1α is independent of the ESCC miRNAs. Interestingly, studies have shown that HIF1α is a direct target of miR-17-92 cluster of miRNAs, which are highly expressed in ESCs and have a similar seed sequence to the ESCC miRNAs of 'AAAGUG', therefore share targets with the ESCC miRNAs such as cdkn1a (148,275). Thus the loss of miRNAs particularly, the miR-17-92 cluster may account for the significantly higher expression of HIF1α in the DGCR8$^{-/-}$ mESCs, and may potentially explain the loss of the link between the expression of HIF1α

and the glucose metabolism enzymes (70). This awaits further experimental confirmation.

Interestingly, a high expression of HIF1α has been implicated in the induction of glutaminolysis in cancer cells (249) and HIF1α was shown to be consistently expressed at higher levels in the DGCR8$^{-/-}$ mESCs compared to the wild type in this study.

10.4.2 Glutamic pyruvate tranaminase 2 (Gpt2) and Isocitrate dehydrogenase 1 (Idh1) are both indirectly regulated by the ESCC miRNAs

In order to verify whether there was a switch from glucose metabolism to glutaminolysis in the DGCR8$^{-/-}$ cells for energy production, the gene expression of glutamic pyruvate transaminase 2 (*gpt2)*, the key enzyme which catalyses the reversible conversion of glutamate to alanine and α-ketoglutarate was analysed.

Gpt2 showed significantly higher expression in the DGCR8$^{-/-}$ mESCs compared to the WT, and this was enhanced by the overexpression of the ESCC miRNAs. This was also the case with the cytosolic Isocitrate Dehydrogenase 1, which showed similar expression at baseline between the two cell types, but was indirectly targeted by the ESCC miRNAs resulting in expression levels that were above WT levels. Notably, both Idh1 and Gpt2 catalyse irreversible reactions, therefore, the higher expression of Idh1 and Gpt2 in the presence of the ESCC miRNAs coupled to the lower intracellular concentration of alanine and higher intracellular concentration of lactate indicate that the presence of the ESCC miRNAs may push mESC metabolism towards glycolysis and in the absence of these miRNAs, the cells utilize more glutaminolysis. This is supported by studies which show that in cancer cells the glutaminolysis pathway is reliant on Idh1 expression (250).

10.4.3 Threonine is secreted at higher levels by the transfected DGCR8$^{-/-}$ than those transfected with the control miRNAs

Threonine is an essential amino acid for mESC pluripotency and survival and is a precursor for glycine synthesis, which in turn is a precursor for purine biosynthesis (114). mESCs have a distinctively high rate of threonine catabolism into glycine and acetyl Co-A, of which the latter can feed into the TCA cycle, providing an alternative energy source for mESCs (111). Interestingly, threonine displayed increased secretion upon the transfection of the individual ESCC miRNAs, with miR-295 causing a significantly higher secretion compared to the DGCR8$^{-/-}$ mESCs transfected with the control miRNA (data not shown).

The lower secretion of threonine into the medium in the presence of the ESCC miRNAs may be indicative of more threonine retention within the transfected DGCR8$^{-/-}$ and may provide a metabolic mechanism for the differentiation defects observed in the DGCR8$^{-/-}$ mESCs. However, as the focus of this study was on mESCs under proliferative conditions, this would require further exploration. Studies have shown that high threonine dehydrogenase activity promotes cellular pluripotency and low activity promotes differentiation (111), so it would be interesting to study the DGCR8$^{-/-}$ under differentiative conditions in this context. The increased retention of threonine in the presence of the ESCC miRNAs, also supports a higher rate of glycolysis in the transfected DGCR8$^{-/-}$ and WT mESCs, as less catabolism of threonine would result in a lower activity of the TCA cycle.

There was no difference in the mRNA expression level of threonine dehydrogenase 2 at baseline between the two cell types, or upon the transfection of the ESCC miRNAs into the DGCR8$^{-/-}$ mESCs. Notably, when Lin28a, a protein which is highly expressed in ESCs and has been successfully been used as part of a cocktail of transcription factors to reprogram human fibroblasts (58), is overexpressed in mESCs, this results in an accumulation of metabolites involved in the threonine-purine biosynthesis pathway (276). Whereas overexpression of the pro-differentiation let-7 family of miRNAs which is usually negatively

regulated in the wild type mESCs by Lin28 proteins (198), reduces the abundance of the metabolites in pathway, indicates that this pathway is regulated by a combination of protein/miRNA interactions that play an important role in the balance between pluripotency and differentiation (276).

Adenylate kinase 2 (closely related to adenylate kinase 4 (Table 7) regulates ATP homeostasis within cells and is therefore linked to purine biosynthesis (277) and was differentially expressed between the WT and DGCR8$^{-/-}$ mESCs at both the protein level and RNA level displaying lower expression in the DGCR8$^{-/-}$ mESCs, which was restored to WT levels in the presence of the ESCC miRNAs further supporting the role of the ESCC miRNAs in purine biosynthesis.

10.5 DGCR8$^{-/-}$ mESCs display higher expression of antioxidant proteins

The comparative DIGE analysis between the DGCR8$^{-/-}$ and WT, with three out of four identified proteins displaying higher expression in the knockout cells.

Of these, Sod1 which showed higher expression in the WT cells compared to the DGCR8$^{-/-}$ in the DIGE analysis, however the at the mRNA level, the opposite pattern of expression was observed. The expression of Sod1 at the RNA level corresponded to the expression pattern observed with the other antioxidant genes; Peroxiredoxin 3 and Thioredoxin-interacting protein 1 (txnip); which both displayed higher expression in the DGCR8$^{-/-}$ mESCs, though the difference observed in txnip gene expression was not significant. Notably, Peroxiredoxin 3 and txnip are predicted targets of the ESCC miRNAs (245), and studies indicate that txnip is an important regulator of cellular glucose and fatty acid metabolism (278), further implicating miRNAs in the regulation of mESC metabolism. The downregulation of txnip has been shown to result in increased rates of glycolysis, as observed in the WT mESCs (278). Txnip's expression was lowered in the presence of the ESCC miRNAs, significantly by miR-294.

SOD1 and Peroxiredoxin 3, are both antioxidant genes, which were expressed at a significantly higher level in the DGCR8$^{-/-}$ cells compared to the WT, and at the qPCR level were shown to be influenced by the expression of the ESCC miRNAs;

SOD1, indirectly by miR-291a, miR-295 and miR-302d, and Peroxiredoxin 3 directly by miR-295. The targeting of Peroxiredoxin 3 by miR-295, would have to be confirmed by luciferase assays.

Of the metabolites which are associated with both the TCA cycle and glycolysis, both alanine and NAD(H) displayed differences upon the transfection of the miRNAs, not only suggesting a possible change in substrate utilization for energy generation, but also changes in the redox status of the cells transfected with the miRNAs. Due to the nature of the analysis NAD^+ and NADH were taken as one measurement, therefore we cannot definitively determine the redox status of the transfected cells, and this is also the case with the measurement of ADP+ATP.

10.6 Chaperone proteins

Overall, the largest groups of proteins that was identified as being differentially expressed in DGCR8$^{-/-}$ mESCs at baseline were chaperone proteins whose major function is to transport and fold proteins post-translation into their final configurations (279). The miRNA walk database (245), which cross references a number of miRNA target prediction databases was used to identify potential targets of the highly expressed miR-295 cluster of miRNAs (48). It revealed two of the chaperone proteins; peptidyl-prolyl cis-trans isomerase FKBP4 and the mitochondrial 60kDa heat shock protein as putative targets of members of the ESCC cluster (241,242,280). The majority of chaperone proteins identified displayed higher expression in the DGCR8$^{-/-}$ compared to the WT mESCs, suggesting that their expression may be regulated by miRNAs. Notably the loss of miRNAs in mESCs has been linked to autophagy associated with an increased expression of chaperone proteins (281).

The common pathway analysis comparing the WT and DGCR8$^{-/-}$ mESCs also supports the finding that the chaperone proteins may be altered as pathways including 'Binding of Erp57' (Erp57 interacts with calreticulin and calnexin to promote their function of folding glycoproteins (282)), 'Activation of chaperones by ATF6-alpha' and 'Protein folding' have been highlighted as being important in

the DGCR8$^{-/-}$ cells. 'Binding of Erp57' even had the highest ratio of proteins identified to be part of that pathway compared to any other pathway, further supporting a role for chaperone proteins in DGCR8$^{-/-}$ mESCs .

Two of the chaperone proteins identified in the DIGE study were validated at the protein level; the expression levels for Serpin H1 were in agreement with the data from the DIGE analysis however, the expression levels for PDI, another chaperone protein, were not, most likely because, PDI is a redox regulated chaperone protein, and therefore has different oxidation states (283). Thus the discrepancy between the differential expression of PDI observed on the DIGE gel and the similar expression of PDI in both the WT and DGCR8$^{-/-}$ cells at the protein level, most likely represents the ability of the DIGE technique to detect different post translational modifications of the same protein, whereas the antibody used in the western blot was not specific for different oxidation states.

Serpin H1, a chaperone protein for collagen, which at the protein level was consistently expressed at a lower level in the DGCR8$^{-/-}$ mESCs compared to the WT, was unaffected by the expression of the ESCCs at the protein level. However ANOVA analysis revealed there to be a highly significant difference (p-value = 0.0021) between the transfected cells at the mRNA level, further t-tests revealed that the mESCs transfected with miR-291b and miR-294 showed significantly higher and lower expression of Serpin H1, respectively compared to the cells transfected with the control miRNA.

Notably, the loss of miRNAs in mESCs has been linked to autophagy coupled with an increased expression of chaperone proteins (281) as observed in this study. Other potentially contributing factors to this trend could be the loss of miRNAs indirectly regulating chaperone protein expression, or resulting in a need for higher chaperone protein expression to deal with the increased number of proteins being translated such as is observed during the unfolded protein response (UPR) (279). In order to test this, the expression of master regulators of the UPR such as PERK, CHOP and ATF6 was analysed, however the results were inconsistent (data not shown).

11 Limitations of the study and future work

11.1 Cell characterization

The cells were characterised at baseline using PCR, Western blots and morphological analysis (Section 9.1),which revealed that both cell types were undifferentiated. However, the alkaline phosphatase staining revealed that once the cells were kept in culture for longer (5 days), differences were revealed both in terms of morphology as shown by the differences in colony formation and the degree of staining (Figure 13). This may suggest a difference in the differentiation status of the cells , which is supported by the DIGE analysis which reveals a higher expression of markers of early differentiation such as Calponin 3 and Vimentin (Table 9) in the DGCR8$^{-/-}$ mESCs compared to the wild type. In order to verify this, both cell types would have to be grown under the same differentiative conditions. Although the cells were counted immediately prior to harvest, to ensure that a similar number of cells had survived and a cell death assay (Figure 18) was also conducted to verify this, however the cell death assay measured the number of nucleosomes released from the cells so only focused on one type of cell death; apoptosis, at one-time point, therefore it would be useful to carry out assays focusing on other forms of cell death at a range of time points.

11.2 Technical difficulties

Although DIGE analysis provides a robust and accurate method to study and compare the proteome of a number of samples, once the proteins have been identified, they need to be validated in order to ensure that the DIGE analysis is accurate. The proteins were run on a gel over a size range of 15-100kDa and a pI ranging from 3-10, which gives an overview of abundant proteins, but only within this mass and pI range. Using common pathway analysis allows for the identification of networks of proteins that may be affected aiding in the identification of proteins missed by DIGE analysis . Therefore, it was vital to integrate a number of approaches in order to obtain the most accurate and in-depth analysis of the mESC phenotype. This is especially true when studying miRNAs effects on protein expression, as miRNAs can act as 'rheostats' for protein expression, so any subtle differences in protein expression may not be

identified by DIGE analysis. For example, although it is clear that Aldolase A expression is influenced by all of the ESCC miRNAs, the DIGE analysis indicated that miR-295 was the only miRNA that had a significant influence on Aldoa expression, miR-294 had a near to significant effect on Aldoa expression, with a p-value at 0.08, while the remaining miRNAs failed to induce a change that reached statistical significance.

11.2 Future work

To further increase the understanding of the underlying mechanisms of the balance between pluripotency and differentiation and to fully study the metabolism of ESCs, it would be imperative to study the cells over a timecourse.

Based on the metabolic differences observed between the DGCR8$^{-/-}$ and WT mESCs, with the DGCR8$^{-/-}$ possibly using glutaminolysis for their energy generation. An important next step would be to provide WT mESCs with glutamine as their major substrate and identify what effect this would have their balance between pluripotency and differentiation. However due to the mESCs' obvious reliance on glucose metabolism for energy production as well as the formation of amino acids and metabolites for cell growth and other intracellular processes, it can be difficult to study any direct effects of altering metabolism which would of course be tightly intertwined with cell survival.

As aforementioned the DGCR8$^{-/-}$ mESCs provide a relatively 'clean' background for the study of miRNA function and targets (252). Thus, studying a highly related set of miRNAs, which are known to have a large number of redundant targets (8,26), does not provide a full representation of the large number of interactions occurring in mESCs. This study focused on the miR-295 cluster and miR-302d due to their highly conserved nature, and universality in ESCs. It has been established that although these miRNAs can be used to reprogram cells, when transfected into DGCR8$^{-/-}$ mESCs they are only able to restore the cell cycle and not any of the differentiation defects. Therefore it would be useful to study

other miRNAs such as the let-7 miRNAs, which are known to have the ability to downregulate the pluripotency markers in DGCR8$^{-/-}$ mESCs (198), either on their own or in conjunction with the ESCC miRNAs.

mESCs are continually poised between differentiation and pluripotency, therefore to fully understand the mechanisms controlling the obvious links between pluripotency, metabolism and miRNA regulation, it would be imperative to conduct the same study but under differentiation conditions. However, this would introduce more questions, as in culture of mESCs, although the majority of cells are undergoing self renewal, a small percentage of cells will be spontaneously, randomly differentiating, therefore the mere removal of LIF would result in a differentiation into a variety of cell types, which would confound any downstream metabolic and proteomic study. Thus a method of differentiation such as growing the mESCs on a collagen IV matrix in the absence of LIF, which would force the majority of cells to differentiate towards the smooth muscle cell lineage (284), would provide a more coherent overview of differentiation, albeit in relation to one lineage. The spontaneous differentiation of mESCs is a major reason for using mESCs of early passages for any experiments, particularly those focused on pluripotency.

12. Conclusion

Increasingly, metabolism and glucose metabolism in particular have been shown to play a direct role in the regulation of the onset of differentiation (29,30,60,63,67). Using a range of combined –omics techniques, we were able to show that DGCR8$^{-/-}$ mESCs which are known to have impaired differentiation (6), not only display lower expression of glycolytic enzymes, but also have a lower glycolytic function than their WT counterparts, which may at least in part account for the differentiation defects.

Upon the re-introduction of the ESCC miRNAs into the DGCR8$^{-/-}$ mESCs, not only was the expression of the glycolytic genes restored to WT levels, most strikingly for Aldoa, but NMR analysis showed that the DGCR8$^{-/-}$ cells secreted more lactate, indicative of a higher rate of glycolysis. Therefore the same miRNAs that control the ESC cycle also influence the metabolic status of the ESCs. Of the miRNAs studied, miR-295 had the most indirect/direct targets in the DIGE studies, and had the greatest effect upon the glucose metabolism enzymes. The lower glycolytic activity of the DGCR8$^{-/-}$ mESCs compared to the WT indicate that the DGCR8$^{-/-}$ may be using a compensatory mechanism for energy generation. Using metabolomics, we discovered potential pathways that the DGCR8$^{-/-}$ mESCs may be using and formulated a putative model for DGCR8$^{-/-}$ cell metabolism, using glutamine as their main source of energy production as opposed to glucose .

The proposed method of energy generation in both cell types in summarized in Figure 42.

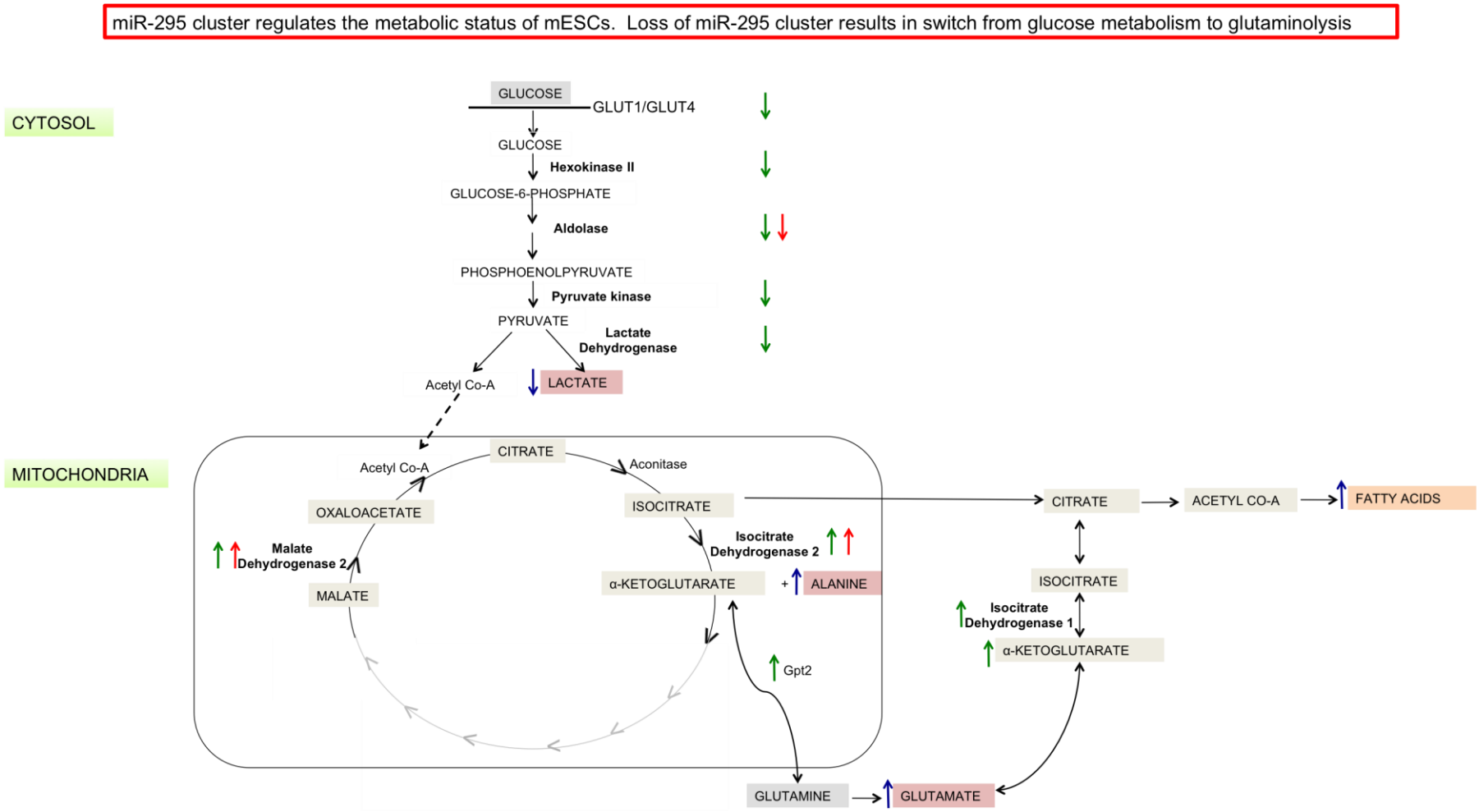


Figure 42- Summary figure of possible glutaminolysis in DGCR8$^{-/-}$ mESCs. Red arrows represent protein expression, green arrows represent mRNA expression and blue arrows represent metabolite changes in DGCR8$^{-/-}$ mESCs in relation to wild type

13. References

(1) Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. Nature 2014 Jan 16;505(7483):344-352.
(2) Sayed D, Abdellatif M. MicroRNAs in development and disease. Physiol Rev 2011 Jul;91(3):827-887.
(3) Wang Y, Baskerville S, Shenoy A, Babiarz JE, Baehner L, Blelloch R. Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. Nature Genetics 2008;40(12):1478-1482.
(4) Lenkala D, LaCroix B, Gamazon ER, Geeleher P, Im HK, Huang RS. The impact of microRNA expression on cellular proliferation. Hum Genet 2014 Jul;133(7):931-938.
(5) Leung AK, Sharp PA. MicroRNA functions in stress responses. Mol Cell 2010 Oct 22;40(2):205-215.
(6) Wang Y, Medvid R, Melton C, Janenisch R, Blelloch R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell-self renewal. Nature Genetics 2007;39(3):116-125.
(7) Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, et al. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev 2005 Feb 15;19(4):489-501.
(8) Davis MP, Abreu-Goodger C, van Dongen S, Lu D, Tate PH, Bartonicek N, et al. Large-scale identification of microRNA targets in murine Dgcr8-deficient embryonic stem cell lines. PLoS One 2012;7(8):e41762.
(9) Mitalipov S, Wolf D. Totipotency, pluripotency and nuclear reprogramming. Adv Biochem Eng Biotechnol 2009;114:185-199.
(10) Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature 1981 Jul 9;292(5819):154-156.
(11) Evans M. Origin of mouse embryonal carcinoma cells and the possibility of their direct isolation into tissue culture. J Reprod Fertil 1981 Jul;62(2):625-631.
(12) Kerr CL, Shamblott MJ, Gearhart JD. Pluripotent stem cells from germ cells. Methods Enzymol 2006;419:400-426.
(13) Andrews PW, Matin MM, Bahrami AR, Damjanov I, Gokhale P, Draper JS. Embryonic stem (ES) cells and embryonal carcinoma (EC) cells: opposite sides of the same coin. Biochem Soc Trans 2005 Dec;33(Pt 6):1526-1530.
(14) Turnpenny L, Brickwood S, Spalluto CM, Piper K, Cameron IT, Wilson DI, et al. Derivation of human embryonic germ cells: an alternative source of pluripotent stem cells. Stem Cells 2003;21(5):598-609.
(15) Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007 Nov 30;131(5):861-872.
(16) Anokye-Danso F, Trivedi CM, Juhr D, Gupta M, Cui Z, Tian Y, et al. Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. Cell Stem Cell 2011 Apr 8;8(4):376-388.
(17) Hou P, Li Y, Zhang X, Liu C, Guan J, Li H, et al. Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. Science 2013 Aug 9;341(6146):651-654.

(18) Narsinh KH, Plews J, Wu JC. Comparison of human induced pluripotent and embryonic stem cells: fraternal or identical twins? Mol Ther 2011 Apr;19(4):635-638.
(19) Stadtfeld M, Apostolou E, Akutsu H, Fukuda A, Follett P, Natesan S, et al. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. Nature 2010 May 13;465(7295):175-181.
(20) Longo L, Bygrave A, Grosveld FG, Pandolfi PP. The chromosome make-up of mouse embryonic stem cells is predictive of somatic and germ cell chimaerism. Transgenic Research 1997;6(5):321-328.
(21) Gangaraju VK, Lin H. MicroRNAs: key regulators of stem cells. Nature Reviews Molecular Cell Biology 2009;10:116-125.
(22) Hiyama E, Hiyama K. Telomere and telomerase in stem cells. Br J Cancer 2007 Apr 10;96(7):1020-1024.
(23) Stead E, White J, Faast R, Conn S, Goldstone S, Rathjen J, et al. Pluripotent cell division cycles are driven by ectopic Cdk2, cyclin A/E and E2F activites. Oncogene 2002;21(54):8320-8333.
(24) Welham M, Kingham E, Sanchez-Ripoll Y, Kumpfmueller B, Storm M, Bone H. Controlling embryonic stem cell proliferation and pluripotency: the role of PI3K- and GSK-3-dependent signalling. Biochemical Society Transactions 2011;39(2):674-678.
(25) Hu B, Mitra J, van den Heuvel S, Enders GH. S and G2 phase roles for Cdk2 revealed by inducible expression of a dominant-negative mutant in human cells. Mol Cell Biol 2001 Apr;21(8):2755-2766.
(26) Tiscornia G, Izpisua Belmonte JC. MicroRNAs in embryonic stem cell function and fate. Nature 2010;24(24):2732-2741.
(27) Coronado D, Godet M, Bourillot PY, Tapponnier Y, Bernat A, Petit M, et al. A short G1 phase is an intrinsic determinant of naive embryonic stem cell pluripotency. Stem Cell Res 2013 Jan;10(1):118-131.
(28) Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. Science 2009;324(5930):1029-1033.
(29) Panopoulos AD, Izpisua Belmonte JC. Anaerobicizing into pluripotency. Cell Metab 2011 Aug 3;14(2):143-144.
(30) Li Q, Hakimi P, Liu X, Yu WM, Ye F, Fujioka H, et al. Cited2, a transcriptional modulator protein, regulates metabolism in murine embryonic stem cells. J Biol Chem 2014 Jan 3;289(1):251-263.
(31) Evan MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature 1981;292:154-156.
(32) Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarinoma cells. Proceedings of the National Academy of Sciences 1981;78:7634-7638.
(33) Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, et al. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature 1988;336:684-687.
(34) Burdon T, Stacey C, Chambers I, Nichols J, Smith A. Suppresion of SHP-2 and ERK signalling promotes self renewal of mouse embryonic stem cells. Developmental Biology 1999;210:30-43.
(35) Ng H, Surani MA. The transcriptional and signalling networks of pluripotency. Naure Cell Biology 2011;13(5):490-496.

(36) Wobus AM, Boheler KR. Embryonic stem cells: prospects for developmental biology and cell therapy. Physiol Rev 2005 Apr;85(2):635-678.
(37) Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nature Genetics 2006;38(4):431-440.
(38) Ying Q, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, et al. The ground state of embryonic stem cell self renewal. Nature 2008;453:519-523.
(39) Avilion AA, Nicolis SK, Peny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell linages in early mouse development depend on SOX2 function. Genes and Development 2003;17:126-140.
(40) Li Y, Zhang Q, Yin X, Yang W, Du Y, Hou P, et al. Generation of iPSCs from mouse fibroblasts with a single gene, Oct4 and small molecules. Cell Research 2011;21(1):196-204.
(41) Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006;126(5):636-676.
(42) Zhu S, Li W, Zhou H, Wei W, Ambasudhan R, Lin T, et al. Reprogramming of human primary somatic cells by OCT4 and chemical compounds. Cell Stem Cell 2010 Dec 3;7(6):651-655.
(43) Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006 Aug 25;126(4):663-676.
(44) Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, et al. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol 2008 Jan;26(1):101-106.
(45) Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nature Genetics 2000;24:372-376.
(46) Matoba R, Niwa H, Masui S, Ohtsuka S, Carter MG, Sharov AA, et al. Dissecting Oct3/4-regulated gene networks in embryonic stem cells by expression profiling. PLoS One 2006 Dec 20;1:e26.
(47) Chew JL, Loh YH, Zhang W, Chen X, Tam WL, Yeap LS, et al. Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. Molecular Cell Biology 2005;25(14):6031-6046.
(48) Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, Johnstone S, et al. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell 2008 Aug 8;134(3):521-533.
(49) Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, et al. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nature Cell Biology 2007;9:625-635.
(50) Yamaguchi S, Hirano K, Nagata S, Tada T. Sox2 expression effects on direct reprogramming efficiency as determined by alternative somatic cell fate. Stem Cell Research 2011 3;6(2):177-186.
(51) Glauche I, Herberg M, Roeder I. Nanog Variability and Pluripotency Regulation of Embryonic Stem Cells - Insights from a Mathematical Model Analysis. PLoS ONE 2010;5(6).
(52) Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev 2003 Jan 1;17(1):126-140.

(53) Ferri AL, Cavallaro M, Braida D, Di Cristofano A, Canta A, Vezzani A, et al. Sox2 deficiency causes neurodegeneration and impaired neurogenesis in the adult mouse brain. Development 2004 Aug;131(15):3805-3819.
(54) Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Kazutoshi T, et al. The Homeoprotein Nanog Is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells. Cell 2003;113(5).
(55) Moon J, Yun W, Kim J, Hyeon S, Kang PJ, Park G, et al. Reprogramming of mouse fibroblasts into induced pluripotent stem cells with Nanog. Biochem Biophys Res Commun 2013 2/15;431(3):444-449.
(56) Chambers I, Tomlinson SR. The transcriptional foundation of pluripotency. Development 2009;136:2311-2322.
(57) Buganim Y, Faddah DA, Jaenisch R. Mechanisms and models of somatic cell reprogramming. Nat Rev Genet 2013 Jun;14(6):427-439.
(58) Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science 2007 Dec 21;318(5858):1917-1920.
(59) Hanna LA, Foreman RK, Tarasenko IA, Kessler DS, Labosky PA. Requirement of Foxd3 in maintaining pluripotent cells of the early mouse embryo. Genes and Development 2002;16:2650-2661.
(60) Yu W, Liu X, Shen J, Jovanovic O, Pohl EE, Gerson S, et al. Metabolic Regulation by the Mitochondrial Phosphatase PTPMT1 Is Required for Hematopoietic Stem Cell Differentiation. Cell Stem Cell 2013;12:62-74.
(61) Varum S, Momcilovic O, Castro C, Ben-Yehudah A, Ramalho-Santos J, Navara CS. Enhancement of human embryonic stem cell pluripotency through inhibition of the mitochondrial respiratory chain. Stem Cell Res 2009 Sep-Nov;3(2-3):142-156.
(62) Zhang J, Nuebel E, Daley G, Koehler C, Teitell M. Metabolic Regulation in Pluripotent Stem Cells during Reprogramming and Self-Renewal. Cell Stem Cell 2012 11/2;11(5):589-595.
(63) Folmes CDL, Nelson TJ, Martinez-Fernandez A, Arrell DK, Lindor JZ, Dzeja PP, et al. Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. Cell Metabolism 2011;14(2):264-271.
(64) Tonack S, Rolletschek A, Wobus AM, Fischer B, Santos AN. Differential expression of glucose transporter isoforms during embryonic stem cell differentiation. Differentiation 2006 Dec;74(9-10):499-509.
(65) Wood IS, Trayhurn P. Glucose transporters (GLUT and SGLT): expanded families of sugar transport proteins. Br J Nutr 2003 Jan;89(1):3-9.
(66) Berg JM, Tymoczko JL, Stryer L. Many shuttles allow Movement across the Mitochondrial membrane. Biochemistry New York: W. H. Freeman and Company; 2002.
(67) Folmes CD, Dzeja PP, Nelson TJ, Terzic A. Metabolic plasticity in stem cell homeostasis and differentiation. Cell Stem Cell 2012 Nov 2;11(5):596-606.
(68) Berg JM, Tymoczko JL, Stryer L. The Glycolytic Pathway Is Tightly Controlled. In: Freeman WH, editor. Biochemistry New York; 2002.
(69) Warburg O. On the origin of cancer cells. Science 1956 Feb 24;123(3191):309-314.

(70) Zhou W, Choi M, Margineantu D, Margaretha L, Hesson J, Cavanaugh C, et al. HIF1alpha induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. EMBO J 2012 May 2;31(9):2103-2116.
(71) Kim YH, Heo JS, Han HJ. High glucose increase cell cycle regulatory proteins level of mouse embryonic stem cells via PI3-K/Akt and MAPKs signal pathways. J Cell Physiol 2006 Oct;209(1):94-102.
(72) Vogt C, Ardehali H, Iozzo P, Yki-Jarvinen H, Koval J, Maezono K, et al. Regulation of hexokinase II expression in human skeletal muscle in vivo. Metabolism 2000 Jun;49(6):814-818.
(73) Sharma B. Kinetic Characterisation of Phosphofructokinase Purified from Setaria cervi: A Bovine Filarial Parasite. Enzyme Res 2011;2011:939472.
(74) Spoden GA, Rostek U, Lechner S, Mitterberger M, Mazurek S, Zwerschke W. Pyruvate kinase isoenzyme M2 is a glycolytic sensor differentially regulating cell proliferation, cell size and apoptotic cell death dependent on glucose supply. Exp Cell Res 2009 10/1;315(16):2765-2774.
(75) Christofk HR, Vander Heiden MG, Harris MH, Ramanathan A, Gerszten RE, Wei R, et al. The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. Nature 2008 Mar 13;452(7184):230-233.
(76) Bluemlein K, Gruning NM, Feichtinger RG, Lehrach H, Kofler B, Ralser M. No evidence for a shift in pyruvate kinase PKM1 to PKM2 expression during tumorigenesis. Oncotarget 2011 May;2(5):393-400.
(77) Dombrauckas JD, Santarsiero BD, Mesecar AD. Structural basis for tumor pyruvate kinase M2 allosteric regulation and catalysis. Biochemistry 2005;44(27):9417-9429.
(78) Mazurek S, Boschek CB, Hugo F, Eigenbrodt E. Pyruvate kinase type M2 and its role in tumor growth and spreading. Semin Cancer Biol 2005 8;15(4):300-308.
(79) Noguchi T, Inoue H, Tanaka T. The M1- and M2-type isozymes of rat pyruvate kinase are produced from the same gene by alternative RNA splicing. J Biol Chem 1986 Oct 15;261(29):13807-13812.
(80) Wong N, De Melo J, Tang D. PKM2, a Central Point of Regulation in Cancer Metabolism. Int J Cell Biol 2013;2013:242513.
(81) Ichai C, El-Mir MY, Nogueira V, Piquet MA, Chauvin C, Fontaine E, et al. Exogenous Mg-ATP induces a large inhibition of pyruvate kinase in intact rat hepatocytes. J Biol Chem 2001 Mar 2;276(9):6398-6403.
(82) Christofk HR, Vander Heiden MG, Wu N, Asara JM, Cantley LC. Pyruvate kinase M2 is a phosphotyrosine-binding protein. Nature 2008 Mar 13;452(7184):181-186.
(83) Mazurek S. Effect of the tetrameric and dimeric form of M2-PK on cell metabolism. 2011; Available at: http://www.metabolic-database.com/html/m2-pk_tetramer_dimer.html. Accessed March/20, 2015.
(84) Gupta V, Bamezai RN. Human pyruvate kinase M2: a multifunctional protein. Protein Sci 2010 Nov;19(11):2031-2044.
(85) Lee J, Kim HK, Han YM, Kim J. Pyruvate kinase isozyme type M2 (PKM2) interacts and cooperates with Oct-4 in regulating transcription. Int J Biochem Cell Biol 2008;40(5):1043-1054.
(86) Yang W, Xia Y, Ji H, Zheng Y, Liang J, Huang W, et al. Nuclear PKM2 regulates beta-catenin transactivation upon EGFR activation. Nature 2011 Dec 1;480(7375):118-122.

(87) Folmes CDL, Nelson TJ, Martinez-Fernandez A, Kent Arrell D, Lindon JZ, Dzeja PP, et al. Somatic Oxidative Bioenergetics Transitions into Pluripotency-Dependent Glycolysis to Facilitate Nuclear Reprogramming. Cell Metabolism 2011;14:264-271.
(88) Panopoulos AD, Yanes O, Ruiz S, Kida YS, Diep D, Tautenhahn R, et al. The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming. Cell Research 2012;22:168-177.
(89) Xu X, Duan S, Yi F, Ocampo A, Liu GH, Izpisua Belmonte JC. Mitochondrial regulation in pluripotent stem cells. Cell Metab 2013 Sep 3;18(3):325-332.
(90) Zhang J, Khvorostov I, Hong JS, Oktay Y, Vergnes L, Nuebel E, et al. UCP2 regulates energy metabolism and differentiation potential of human pluripotent stem cells. EMBO J 2011 Nov 15;30(24):4860-4873.
(91) Wanet A, Remacle N, Najar M, Sokal E, Arnould T, Najimi M, et al. Mitochondrial remodeling in hepatic differentiation and dedifferentiation. Int J Biochem Cell Biol 2014 9;54(0):174-185.
(92) Yu WM, Liu X, Shen J, Jovanovic O, Pohl EE, Gerson SL, et al. Metabolic regulation by the mitochondrial phosphatase PTPMT1 is required for hematopoietic stem cell differentiation. Cell Stem Cell 2013 Jan 3;12(1):62-74.
(93) Rupprecht A, Sittner D, Smorodchenko A, Hilse KE, Goyn J, Moldzio R, et al. Uncoupling protein 2 and 4 expression pattern during stem cell differentiation provides new insight into their putative function. PLoS One 2014 Feb 11;9(2):e88474.
(94) Birket MJ, Orr AL, Gerencser AA, Madden DT, Vitelli C, Swistowski A, et al. A reduction in ATP demand and mitochondrial activity with neural differentiation of human embryonic stem cells. J Cell Sci 2011 Feb 1;124(Pt 3):348-358.
(95) Todd LR, Damin MN, Gomathinayagam R, Horn SR, Means AR, Sankar U. Growth factor erv1-like modulates Drp1 to preserve mitochondrial dynamics and function in mouse embryonic stem cells. Mol Biol Cell 2010 Apr 1;21(7):1225-1236.
(96) Covello KL, Kehler J, Yu H, Gordan JD, Arsham AM, Hu CJ, et al. HIF-2alpha regulates Oct-4: effects of hypoxia on stem cell function, embryonic development, and tumor growth. Genes Dev 2006 Mar 1;20(5):557-570.
(97) Jeong CH, Lee HJ, Cha JH, Kim JH, Kim KR, Kim JH, et al. Hypoxia-inducible factor-1 alpha inhibits self-renewal of mouse embryonic stem cells in Vitro via negative regulation of the leukemia inhibitory factor-STAT3 pathway. J Biol Chem 2007 May 4;282(18):13672-13679.
(98) Lee SW, Jeong HK, Lee JY, Yang J, Lee EJ, Kim SY, et al. Hypoxic priming of mESCs accelerates vascular-lineage differentiation through HIF1-mediated inverse regulation of Oct4 and VEGF. EMBO Mol Med 2012 Sep;4(9):924-938.
(99) Wang Y, Yang J, Li H, Wang X, Zhu L, Fan M, et al. Hypoxia promotes dopaminergic differentiation of mesenchymal stem cells and shows benefits for transplantation in a rat model of Parkinson's disease. PLoS One 2013;8(1):e54296.
(100) Heinis M, Simon MT, Ilc K, Mazure NM, Pouyssegur J, Scharfmann R, et al. Oxygen tension regulates pancreatic beta-cell differentiation through hypoxia-inducible factor 1alpha. Diabetes 2010 Mar;59(3):662-669.
(101) Hu CJ, Wang LY, Chodosh LA, Keith B, Simon MC. Differential roles of hypoxia-inducible factor 1alpha (HIF-1alpha) and HIF-2alpha in hypoxic gene regulation. Mol Cell Biol 2003 Dec;23(24):9361-9374.

(102) Wang V, Davis DA, Haque M, Huang LE, Yarchoan R. Differential gene up-regulation by hypoxia-inducible factor-1alpha and hypoxia-inducible factor-2alpha in HEK293T cells. Cancer Res 2005 Apr 15;65(8):3299-3306.
(103) Yoshida Y, Takahashi K, Okita K, Ichisaka T, Yamanaka S. Hypoxia Enhances the Generation of Induced Pluripotent Stem Cells. Cell Stem Cell 2009 9/4;5(3):237-241.
(104) Becker KA, Stein JL, Lian JB, van Wijnen AJ, Stein GS. Human embryonic stem cells are pre-mitotically committed to self-renewal and acquire a lengthened G1 phase upon lineage programming. J Cell Physiol 2010 Jan;222(1):103-110.
(105) Wang Y, Melton C, Li YP, Shenoy A, Zhang XX, Subramanyam D, et al. miR-294/miR-302 promotes proliferation, suppresses G1-S restriction point, and inhibits ESC differentiation through separable mechanisms. Cell Rep 2013 Jul 11;4(1):99-109.
(106) Heo JS, Han HJ. ATP stimulates mouse embryonic stem cell proliferation via protein kinase C, phosphatidylinositol 3-kinase/Akt, and mitogen-activated protein kinase signaling pathways. Stem Cells 2006 Dec;24(12):2637-2648.
(107) Yanes O, Clark J, Wong DM, Patti GJ, Sanchez-Ruiz A, Benton HP, et al. Metabolic oxidation regulates embryonic stem cell differentiation. Nat Chem Biol 2010 Jun;6(6):411-417.
(108) Kim MH, Kim MO, Kim YH, Kim JS, Han HJ. Linoleic acid induces mouse embryonic stem cell proliferation via Ca2+/PKC, PI3K/Akt, and MAPKs. Cell Physiol Biochem 2009;23(1-3):53-64.
(109) Zaugg K, Yao Y, Reilly PT, Kannan K, Kiarash R, Mason J, et al. Carnitine palmitoyltransferase 1C promotes cell survival and tumor growth under conditions of metabolic stress. Genes Dev 2011 May 15;25(10):1041-1051.
(110) Emre Y, Nübel T. Uncoupling protein UCP2: When mitochondrial activity meets immunity. FEBS Lett 2010 4/16;584(8):1437-1442.
(111) Wang J, Alexander P, Wu L, Hammer R, Cleaver O, McKnight SL. Dependence of mouse embryonic stem cells on threonine catabolism. Science 2009 Jul 24;325(5939):435-439.
(112) Shiraki N, Shiraki Y, Tsuyama T, Obata F, Miura M, Nagae G, et al. Methionine Metabolism Regulates Maintenance and Differentiation of Human Pluripotent Stem Cells. Cell Metabolism 2014 5/6;19(5):780-794.
(113) Edgar AJ. The human L-threonine 3-dehydrogenase gene is an expressed pseudogene. BMC Genet 2002 Oct 2;3:18.
(114) Han C, Gu H, Wang J, Lu W, Mei Y, Wu M. Regulation of L-threonine dehydrogenase in somatic cell reprogramming. Stem Cells 2013 May;31(5):953-965.
(115) Colombo SL, Palacios-Callender M, Frakich N, De Leon J, Schmitt CA, Boorn L, et al. Anaphase-promoting complex/cyclosome-Cdh1 coordinates glycolysis and glutaminolysis with transition to S phase in human T lymphocytes. Proc Natl Acad Sci U S A 2010 Nov 2;107(44):18868-18873.
(116) Zhu H, Shyh-Chang N, Segrè A, Shinoda G, Shah S, Einhorn W, et al. The Lin28/let-7 Axis Regulates Glucose Metabolism. Cell 2011 9/30;147(1):81-94.
(117) Vasudevan S, Tong Y, Steitz JA. Switching from Repression to Activation: MicroRNAs Can Up-Regulate Translation. Science 2007 December 21;318(5858):1931-1934.
(118) Vazquez F. Small RNAs in Plants. eLS: John Wiley & Sons, Ltd; 2001.

(119) Kloosterman WP, Plasterk RH. The diverse functions of microRNAs in animal development and disease. Dev Cell 2006 Oct;11(4):441-450.
(120) Wu L, Fan J, Belasco JG. MicroRNAs direct rapid deadenylation of mRNA. Proc Natl Acad Sci U S A 2006 Mar 14;103(11):4034-4039.
(121) Chekulaeva M, Filipowicz W. Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. Curr Opin Cell Biol 2009 6;21(3):452-460.
(122) Bernstein E, Kim SY, Carmell MA, Murchison EP, Alcorn H, Li MZ, et al. Dicer is essential for mouse development. Nat Genet 2003 Nov;35(3):215-217.
(123) Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 1993 Dec 3;75(5):843-854.
(124) Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, et al. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 2000 Feb 24;403(6772):901-906.
(125) Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, et al. miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A 2005 Sep 27;102(39):13944-13949.
(126) Miyaki S, Sato T, Inoue A, Otsuki S, Ito Y, Yokoyama S, et al. MicroRNA-140 plays dual roles in both cartilage development and homeostasis. Genes Dev 2010 Jun 1;24(11):1173-1185.
(127) Diederichs S, Haber DA. Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. Cell 2007 Dec 14;131(6):1097-1108.
(128) Schanen BC, Li X. Transcriptional regulation of mammalian miRNA genes. Genomics 2011 Jan;97(1):1-6.
(129) Sohn SY, Bae WJ, Kim JJ, Yeom KH, Kim VN, Cho Y. Crystal structure of human DGCR8 core. Nat Struct Mol Biol 2007 Sep;14(9):847-853.
(130) Hutvagner G. Small RNA asymmetry in RNAi: function in RISC assembly and gene regulation. FEBS Lett 2005 Oct 31;579(26):5850-5857.
(131) Guo L, Lu Z. The Fate of miRNA* Strand through Evolutionary Analysis: Implication for Degradation As Merely Carrier Strand or Potential Regulatory Molecule? PLoS ONE 2010;5.
(132) Kuchenbauer F, Mah SM, Heuser M, McPherson A, Ruschmann J, Rouhi A, et al. Comprehensive analysis of mammalian miRNA* species and their role in myeloid cells. Blood 2011 Sep 22;118(12):3350-3358.
(133) Loinger A, Shemla Y, Simon I, Margalit H, Biham O. Competition between small RNAs: a quantitative view. Biophys J 2012 Apr 18;102(8):1712-1721.
(134) Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A. Identification of mammalian microRNA host genes and transcription units. Genome Res 2004 Oct;14(10A):1902-1910.
(135) Li SC, Tang P, Lin WC. Intronic microRNA: discovery and biological implications. DNA Cell Biol 2007 Apr;26(4):195-207.
(136) Kim YK, Kim VN. Processing of intronic microRNAs. EMBO J 2007 Feb 7;26(3):775-783.
(137) Ying SY, Lin SL. Current perspectives in intronic micro RNAs (miRNAs). J Biomed Sci 2006 Jan;13(1):5-15.

(138) Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, et al. The Microprocessor complex mediates the genesis of microRNAs. Nature 2004 Nov 11;432(7014):235-240.
(139) Ruby JG, Jan CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. Nature 2007 Jul 5;448(7149):83-86.
(140) Havens MA, Reich AA, Duelli DM, Hastings ML. Biogenesis of mammalian microRNAs by a non-canonical processing pathway. Nucleic Acids Res 2012 May;40(10):4626-4640.
(141) Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. Genes Dev 2008 Oct 15;22(20):2773-2785.
(142) Pederson T. Regulatory RNAs derived from transfer RNA? RNA 2010 Oct;16(10):1865-1869.
(143) Maute RL, Schneider C, Sumazin P, Holmes A, Califano A, Basso K, et al. tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. Proc Natl Acad Sci U S A 2013 Jan 22;110(4):1404-1409.
(144) Davis MP, Abreu-Goodger C, van Dongen S, Lu D, Tate PH, Bartonicek N, et al. Large-scale identification of microRNA targets in murine Dgcr8-deficient embryonic stem cell lines. PLoS One 2012;7(8):e41762.
(145) Leung AK, Young AG, Bhutkar A, Zheng GX, Bosson AD, Nielsen CB, et al. Genome-wide identification of Ago2 binding sites from mouse embryonic stem cells with and without mature microRNAs. Nat Struct Mol Biol 2011 Feb;18(2):237-244.
(146) Gruber AJ, Grandy WA, Balwierz PJ, Dimitrova YA, Pachkov M, Ciaudo C, et al. Embryonic stem cell-specific microRNAs contribute to pluripotency by inhibiting regulators of multiple differentiation pathways. Nucleic Acids Res 2014;42(14):9313-9326.
(147) Sinkkonen L, Hugenschmidt T, Berninger P, Gaidatzis D, Mohn F, Artus-Revel CG, et al. MicroRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. Nat Struct Mol Biol 2008 Mar;15(3):259-267.
(148) Zheng GX, Ravi A, Calabrese JM, Medeiros LA, Kirak O, Dennis LM, et al. A latent pro-survival function for the mir-290-295 cluster in mouse embryonic stem cells. PLoS Genet 2011 May;7(5):e1002054.
(149) Davis MP, Abreu-Goodger C, van Dongen S, Lu D, Tate PH, Bartonicek N, et al. Large-scale identification of microRNA targets in murine Dgcr8-deficient embryonic stem cell lines. PLoS One 2012;7(8):e41762.
(150) Grosswendt S, Filipchyk A, Manzano M, Klironomos F, Schilling M, Herzog M, et al. Unambiguous Identification of miRNA:Target Site Interactions by Different Types of Ligation Reactions. Mol Cell 2014 6/19;54(6):1042-1054.
(151) Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. Genes & Development 2006 March 01;20(5):515-524.
(152) Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet 2008 Feb;9(2):102-114.
(153) Martinez J, Tuschl T. RISC is a 5' phosphomonoester-producing RNA endonuclease. Genes Dev 2004 May 1;18(9):975-980.

(154) Brennecke J, Stark A, Russell RB, Cohen SM. Principles of microRNA-target recognition. PLoS Biol 2005 Mar;3(3):e85.
(155) Peters L, Meister G. Argonaute Proteins: Mediators of RNA Silencing. Mol Cell 2007 6/8;26(5):611-623.
(156) Kiriakidou M, Tan GS, Lamprinaki S, De Planell-Saguer M, Nelson PT, Mourelatos Z. An mRNA m7G cap binding-like motif within human Ago2 represses translation. Cell 2007 Jun 15;129(6):1141-1151.
(157) Wakiyama M, Takimoto K, Ohara O, Yokoyama S. Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. Genes & Development 2007;21:1857-1862.
(158) Subtelny AO, Eichhorn SW, Chen GR, Sive H, Bartel DP. Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature 2014 Apr 3;508(7494):66-71.
(159) Tang R, Li L, Zhu D, Hou D, Cao T, Gu H, et al. Mouse miRNA-709 directly regulates miRNA-15a/16-1 biogenesis at the posttranscriptional level in the nucleus: evidence for a microRNA hierarchy system. Cell Res 2012 Mar;22(3):504-515.
(160) Haussecker D, Kay MA. miR-122 continues to blaze the trail for microRNA therapeutics. Mol Ther 2010 Feb;18(2):240-242.
(161) Foldes-Papp Z, Konig K, Studier H, Buckle R, Breunig HG, Uchugonova A, et al. Trafficking of mature miRNA-122 into the nucleus of live liver cells. Curr Pharm Biotechnol 2009 Sep;10(6):569-578.
(162) Murchison EP, Partridge JF, Tam OH, Cheloufi S, Hannon GJ. Characterization of Dicer-deficient murine embryonic stem cells. Proceedings of the National Academy of Sciences of the United States of America 2005 August 23;102(34):12135-12140.
(163) Wang P, Bouwman FG, Mariman EC. Generally detected proteins in comparative proteomics--a matter of cellular stress response? Proteomics 2009 Jun;9(11):2955-2966.
(164) Kim BM, Choi MY. Non-canonical microRNAs miR-320 and miR-702 promote proliferation in Dgcr8-deficient embryonic stem cells. Biochem Biophys Res Commun 2012 Sep 21;426(2):183-189.
(165) Wu S, Aksoy M, Shi J, Houbaviy HB. Evolution of the miR-290-295/miR-371-373 cluster family seed repertoire. PLoS One 2014 Sep 30;9(9):e108519.
(166) Bar M, Wyman SK, Fritz BR, Qi J, Garg KS, Parkin RK, et al. MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. Stem Cells 2008 Oct;26(10):2496-2505.
(167) Wilson DI, Burn J, Scrambler P, Goodship J. DiGeorge syndrome:part of CATCH 22. Jounal of Medical Genetics 1993;10:852-856.
(168) Schofield CM, Hsu R, Barker AJ, Gertz CC, Blelloch R, Ullian EM. Monoallelic deletion of the microRNA biogenesis gene Dgcr8 produces deficits in the development of excitatory synaptic transmission in the prefrontal cortex. Neural Dev 2011 Apr 5;6:11-8104-6-11.
(169) Hacihamdioglu B, Berberoglu M, Siklar Z, Dogu F, Bilir P, Savas Erdeve S, et al. Case report: two patients with partial DiGeorge syndrome presenting with attention disorder and learning difficulties. J Clin Res Pediatr Endocrinol 2011;3(2):95-97.
(170) Bushati N, Cohen SM. microRNA functions. Annu Rev Cell Dev Biol 2007;23:175-205.

(171) Sellier C, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, et al. Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in fragile X-associated tremor/ataxia syndrome. Cell Rep 2013 Mar 28;3(3):869-880.
(172) Yeom KH, Lee Y, Han J, Suh MR, Kim VN. Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. Nucleic Acids Res 2006;34(16):4622-4629.
(173) Roth BM, Ishimaru D, Hennig M. The core microprocessor component DiGeorge syndrome critical region 8 (DGCR8) is a nonspecific RNA-binding protein. J Biol Chem 2013 Sep 13;288(37):26785-26799.
(174) Quick-Cleveland J, Jacob JP, Weitz SH, Shoffner G, Senturia R, Guo F. The DGCR8 RNA-binding heme domain recognizes primary microRNAs by clamping the hairpin. 2014 Jun 26;7(6):1994-2005.
(175) Wada T, Kikuchi J, Furukawa Y. Histone deacetylase 1 enhances microRNA processing via deacetylation of DGCR8. EMBO Rep 2012 Feb 1;13(2):142-149.
(176) Herbert KM, Pimienta G, DeGregorio SJ, Alexandrov A, Steitz JA. Phosphorylation of DGCR8 increases its intracellular stability and induces a progrowth miRNA profile. Cell Rep 2013 Nov 27;5(4):1070-1081.
(177) Macias S, Plass M, Stajuda A, Michlewski G, Eyras E, Caceres JF. DGCR8 HITS-CLIP reveals novel functions for the Microprocessor. Nat Struct Mol Biol 2012 Aug;19(8):760-766.
(178) Chong MM, Zhang G, Cheloufi S, Neubert TA, Hannon GJ, Littman DR. Canonical and alternate functions of the microRNA biogenesis machinery. Genes Dev 2010 Sep 1;24(17):1951-1960.
(179) Zheng GX, Do BT, Webster DE, Khavari PA, Chang HY. Dicer-microRNA-Myc circuit promotes transcription of hundreds of long noncoding RNAs. Nat Struct Mol Biol 2014 Jul;21(7):585-590.
(180) Han J, Pedersen JS, Kwon SC, Belair CD, Kim YK, Yeom KH, et al. Posttranscriptional crossregulation between Drosha and DGCR8. Cell 2009;136:75-84.
(181) Thomson JM, Newman M, Parker JS, Morin-Kensicki EM, Wright T, Hammond SM. Extensive post-transcriptional regulation of microRNAs and its implications for cancer. Genes Dev 2006 Aug 15;20(16):2202-2207.
(182) Suzuki HI, Yamagata K, Sugimoto K, Iwamoto T, Kato S, Miyazono K. Modulation of microRNA processing by p53 . Nature 2009;460:529-533.
(183) Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, et al. The nuclear RNase III Drosha initiates microRNA processing. Nature 2003 Sep 25;425(6956):415-419.
(184) Francia S, Michelini F, Saxena A, Tang D, de Hoon M, Anelli V, et al. Site-specific DICER and DROSHA RNA products control the DNA-damage response. Nature 2012 Aug 9;488(7410):231-235.
(185) Chowdhury D, Choi YE, Brault ME. Charity begins at home: non-coding RNA functions in DNA repair. Nat Rev Mol Cell Biol 2013 Mar;14(3):181-189.
(186) Wu H, Xu H, Miraglia LJ, Crooke ST. Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. J Biol Chem 2000 Nov 24;275(47):36957-36965.
(187) Ganesan G, Rao SM. A novel noncoding RNA processed by Drosha is restricted to nucleus in mouse. RNA 2008 Jul;14(7):1399-1410.

(188) O'Carroll D, Mecklenbrauker I, Das PP, Santana A, Koenig U, Enright AJ, et al. A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. Genes Dev 2007 Aug 15;21(16):1999-2004.
(189) Su H, Trombly MI, Chen J, Wang X. Essential and overlapping functions for mammalian Argonautes in microRNA silencing. Genes Dev 2009 Feb 1;23(3):304-317.
(190) Azuma-Mukai A, Oguri H, Mituyama T, Qian ZR, Asai K, Siomi H, et al. Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. Proc Natl Acad Sci U S A 2008 Jun 10;105(23):7964-7969.
(191) Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, et al. Argonaute2 is the catalytic engine of mammalian RNAi. Science 2004 Sep 3;305(5689):1437-1441.
(192) Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell 2004 Jul 23;15(2):185-197.
(193) Gibbings D, Mostowy S, Jay F, Schwab Y, Cossart P, Voinnet O. Selective autophagy degrades DICER and AGO2 and regulates miRNA activity. Nat Cell Biol 2012 Dec;14(12):1314-1321.
(194) Pratt AJ, MacRae IJ. The RNA-induced silencing complex: a versatile gene-silencing machine. J Biol Chem 2009 Jul 3;284(27):17897-17901.
(195) Rivas FV, Tolia NH, Song JJ, Aragon JP, Liu J, Hannon GJ, et al. Purified Argonaute2 and an siRNA form recombinant human RISC. Nat Struct Mol Biol 2005 Apr;12(4):340-349.
(196) Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, et al. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature 2005 Aug 4;436(7051):740-744.
(197) Ha M, Kim VN. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol 2014 Aug;15(8):509-524.
(198) Melton C, Judson R,L., Blelloch,Robert. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature 2010;463:621-626.
(199) Karginov FV, Cheloufi S, Chong MM, Stark A, Smith AD, Hannon GJ. Diverse endonucleolytic cleavage sites in the mammalian transcriptome depend upon microRNAs, Drosha, and additional nucleases. Mol Cell 2010 Jun 25;38(6):781-788.
(200) Park CY, Choi YS, McManus MT. Analysis of microRNA knockouts in mice. Hum Mol Genet 2010 Oct 15;19(R2):R169-75.
(201) Shyh-Chang N, Daley G. Lin28: Primal Regulator of Growth and Metabolism in Stem Cells. Cell Stem Cell 2013 4/4;12(4):395-406.
(202) Ramirez CM, Goedeke L, Rotllan N, Yoon JH, Cirera-Salinas D, Mattison JA, et al. MicroRNA 33 regulates glucose metabolism. Mol Cell Biol 2013 Aug;33(15):2891-2902.
(203) Viswanathan SR, Daley GQ, Gregory RI. Selective blockade of microRNA processing by Lin28. Science 2008 Apr 4;320(5872):97-100.
(204) Calabrese JM, Seila AC, Yeo GW, Sharp PA. RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. Proc Natl Acad Sci U S A 2007 Nov 13;104(46):18097-18102.
(205) Houbaviy HB, Murray MF, Sharp PA. Embryonic Stem Cell-Specific MicroRNAs. Developmental Cell 2003 8;5(2):351-358.

(206) Suh M, Lee Y, Kim JY, Kim S, Moon S, Lee JY, et al. Human embryonic stem cells express a unique set of microRNAs. Dev Biol 2004 6/15;270(2):488-498.
(207) Choi WY, Giraldez AJ, Schier AF. Target protectors reveal dampening and balancing of Nodal agonist and antagonist by miR-430. Science 2007 Oct 12;318(5848):271-274.
(208) Lichner Z, Pall E, Kerekes A, Pallinger E, Maraghechi P, Bosze Z, et al. The miR-290-295 cluster promotes pluripotency maintenance by regulating cell cycle phase distribution in mouse embryonic stem cells. Differentiation 2011 Jan;81(1):11-24.
(209) Zovoilis A, Smorag L, Pantazi A, Engel W. Members of the miR-290 cluster modulate in vitro differentiation of mouse embryonic stem cells. Differentiation 2009 Sep-Oct;78(2-3):69-78.
(210) Card DA, Hebbar PB, Li L, Trotter KW, Komatsu Y, Mishina Y, et al. Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. Mol Cell Biol 2008 Oct;28(20):6426-6438.
(211) Johnson CD, Esquela-Kerscher A, Stefani G, Byrom M, Kelnar K, Ovcharenko D, et al. The let-7 microRNA represses cell proliferation pathways in human cells. Cancer Res 2007 Aug 15;67(16):7713-7722.
(212) Medeiros LA, Dennis LM, Gill ME, Houbaviy H, Markoulaki S, Fu D, et al. Mir-290-295 deficiency in mice results in partially penetrant embryonic lethality and germ cell defects. Proc Natl Acad Sci U S A 2011 Aug 23;108(34):14163-14168.
(213) Chen Y, Liersch R, Detmar M. The miR-290-295 cluster suppresses autophagic cell death of melanoma cells. Sci Rep 2012;2:808.
(214) Judson RL, Babiarz JE, Venere M, Blelloch R. Embryonic stem cell-specific microRNAs promote induced pluripotency. Nat Biotechnol 2009 May;27(5):459-461.
(215) Kim BM, Thier MC, Oh S, Sherwood R, Kanellopoulou C, Edenhofer F, et al. MicroRNAs are indispensable for reprogramming mouse embryonic fibroblasts into induced stem cell-like cells. PLoS One 2012;7(6):e39239.
(216) Gundry RL, Tchernyshyov I, Sheng S, Tarasova Y, Raginski K, Boheler KR, et al. Expanding the mouse embryonic stem cell proteome: Combining three proteomic approaches. Proteomics 2010;10(14):2728-2732.
(217) Rigbolt KT, Prokhorova TA, Akimov V, Henningsen J, Johansen PT, Kratchmarova I, et al. System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation. Sci Signal 2011 Mar 15;4(164):rs3.
(218) Fathi A, Pakzad M, Taei A, Brink TC, Pirhaji L, Ruiz G, et al. Comparative proteome and transcriptome analyses of embryonic stem cells during embryoid body-based differentiation. Proteomics 2009 Nov;9(21):4859-4870.
(219) Bodzon-Kulakowska A, Bierczynska-Krzysik A, Dylag T, Drabik A, Suder P, Noga M, et al. Methods for samples preparation in proteomic research. J Chromatogr B Analyt Technol Biomed Life Sci 2007 Apr 15;849(1-2):1-31.
(220) Chevalier F. Highlights on the capacities of "Gel-based" proteomics. Proteome Sci 2010 Apr 28;8:23-5956-8-23.
(221) Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. Anal Bioanal Chem 2007 Oct;389(4):1017-1031.

(222) Carrette O, Burkhard PR, Sanchez JC, Hochstrasser DF. State-of-the-art two-dimensional gel electrophoresis: a key tool of proteomics research. Nat Protoc 2006;1(2):812-823.
(223) May C, Brosseron F, Chartowski P, Meyer HE, Marcus K. Differential proteome analysis using 2D-DIGE. Methods Mol Biol 2012;893:75-82.
(224) Rosengren AT, Salmi JM, Aittokallio T, Westerholm J, Lahesmaa R, Nyman TA, et al. Comparison of PDQuest and Progenesis software packages in the analysis of two-dimensional electrophoresis gels. Proteomics 2003 Oct;3(10):1936-1946.
(225) Chevalier F, Rofidal V, Rossignol M. Visible and fluorescent staining of two-dimensional gels. Methods Mol Biol 2007;355:145-156.
(226) Huynh M, Russell P, Walsh B. Tryptic Digestion of In-Gel Proteins for Mass Spectrometric Analysis. In: Tyther R, Sheehan D, editors. Two-Dimensional Electrophoresis Protocols Methods in Molecular Biology Australia: Humana Press; 2009. p. 507-513.
(227) Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem 2001;70:437-473.
(228) Ashcroft AE editor. Ionization Methods in Organic Mass Spectrometry. : Royal Society of Chemistry; 1997.
(229) Ho CS, Lam CW, Chan MH, Cheung RC, Law LK, Lit LC, et al. Electrospray ionisation mass spectrometry: principles and clinical applications. Clin Biochem Rev 2003;24(1):3-12.
(230) Creese AJ, Cooper HJ. Liquid chromatography electron capture dissociation tandem mass spectrometry (LC-ECD-MS/MS) versus liquid chromatography collision-induced dissociation tandem mass spectrometry (LC-CID-MS/MS) for the identification of proteins. J Am Soc Mass Spectrom 2007 May;18(5):891-897.
(231) The UniProt Consortium. Activities at the Universal Protein Resource (UniProt). Nucleic Acids Research 2014 January 01;42(D1):D191-D198.
(232) Granvogl B, Ploscher M, Eichacker LA. Sample preparation by in-gel digestion for mass spectrometry-based proteomics. Anal Bioanal Chem 2007 Oct;389(4):991-1002.
(233) Kruger NJ, Troncoso-Ponce MA, Ratcliffe RG. 1H NMR metabolite fingerprinting and metabolomic analysis of perchloric acid extracts from plant tissues. Nat Protoc 2008;3(6):1001-1012.
(234) Mayr M, Yusuf S, Weir G, Chung YL, Mayr U, Yin X, et al. Combined metabolomic and proteomic analysis of human atrial fibrillation. J Am Coll Cardiol 2008 Feb 5;51(5):585-594.
(235) Geissmann Q. OpenCFU, a New Free and Open-Source Software to Count Cell Colonies and Other Circular Objects. - PLoS ONE (- 2):e54072.
(236) Motulsky H, Christopoulos A. Fiting Models to Biological Data using Linear and Nonlinear Regression. A Practical Guide to Curve Fitting. New York: Oxford University Press; 2004.
(237) Antonov AV. BioProfiling.de: analytical web portal for high-throughput cell biology. Nucleic Acids Res 2011 Jul;39(Web Server issue):W323-7.
(238) Milacic M, Haw R, Rothfels K, Wu G, Croft D, Hermjakob H, et al. Annotating cancer variants and anti-cancer therapeutics in reactome. Cancers (Basel) 2012 Nov 8;4(4):1180-1211.

(239) Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, et al. The Reactome pathway knowledgebase. Nucleic Acids Res 2014 Jan;42(Database issue):D472-7.
(240) Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 2005 Jan 14;120(1):15-20.
(241) Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: targets and expression. Nucleic Acids Research 2008 January 01;36(suppl 1):D149-D153.
(242) Betel D, Koppal A, Agius P, Sander C, Leslie C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. Genome Biol 2010;11(8):R90-2010-11-8-r90. Epub 2010 Aug 27.
(243) Paraskevopoulou MD, Georgakilas G, Kostoulas N, Vlachos IS, Vergoulis T, Reczko M, et al. DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. Nucleic Acids Res 2013 Jul;41(Web Server issue):W169-73.
(244) Reczko M, Maragkakis M, Alexiou P, Grosse I, Hatzigeorgiou AG. Functional microRNA targets in protein coding sequences. Bioinformatics 2012 Mar 15;28(6):771-776.
(245) Dweep H, Sticht C, Pandey P, Gretz N. miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform 2011 Oct;44(5):839-847.
(246) Hsu SD, Chu CH, Tsou AP, Chen SJ, Chen HC, Hsu PW, et al. miRNAMap 2.0: genomic maps of microRNAs in metazoan genomes. Nucleic Acids Res 2008 Jan;36(Database issue):D165-9.
(247) Dennis G,Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003;4(5):P3.
(248) Carey BW, Finley LW, Cross JR, Allis CD, Thompson CB. Intracellular alpha-ketoglutarate maintains the pluripotency of embryonic stem cells. Nature 2014 Dec 10.
(249) Semenza GL. HIF-1 mediates metabolic responses to intratumoral hypoxia and oncogenic mutations. J Clin Invest 2013 Sep;123(9):3664-3671.
(250) Metallo CM, Gameiro PA, Bell EL, Mattaini KR, Yang J, Hiller K, et al. Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. Nature 2011 Nov 20;481(7381):380-384.
(251) Butter F, Scheibe M, Morl M, Mann M. Unbiased RNA-protein interaction screen by quantitative proteomics. Proc Natl Acad Sci U S A 2009 Jun 30;106(26):10626-10631.
(252) Kaspi H, Chapnik E, Levy M, Beck G, Hornstein E, Soen Y. Brief report: miR-290-295 regulate embryonic stem cell differentiation propensities by repressing Pax6. Stem Cells 2013 Oct;31(10):2266-2272.
(253) Sinkkonen L, Hugenschmidt T, Berninger P, Gaidatzis D, Mohn F, Artus-Revel CG, et al. MicroRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. Nat Struct Mol Biol 2008 Mar;15(3):259-267.
(254) Kumar RM, Cahan P, Shalek AK, Satija R, DaleyKeyser AJ, Li H, et al. Deconstructing transcriptional heterogeneity in pluripotent stem cells. Nature 2014 Dec 4;516(7529):56-61.

(255) Guo WT, Wang XW, Yan YL, Li YP, Yin X, Zhang Q, et al. Suppression of epithelial-mesenchymal transition and apoptotic pathways by miR-294/302 family synergistically blocks let-7-induced silencing of self-renewal in embryonic stem cells. Cell Death Differ 2014 Dec 12.
(256) Chakravarti D, Su X, Cho MS, Bui NH, Coarfa C, Venkatanarayan A, et al. Induced multipotency in adult keratinocytes through down-regulation of DeltaNp63 or DGCR8. Proc Natl Acad Sci U S A 2014 Feb 4;111(5):E572-81.
(257) Guo Y, Tian P, Yang C, Liang Z, Li M, Sims M, et al. Silencing the Double-Stranded RNA Binding Protein DGCR8 Inhibits Ovarian Cancer Cell Proliferation, Migration, and Invasion. Pharm Res 2013 Oct 12.
(258) Gomez-Cabello D, Adrados I, Gamarra D, Kobayashi H, Takatsu Y, Takatsu K, et al. DGCR8-mediated disruption of miRNA biogenesis induces cellular senescence in primary fibroblasts. Aging Cell 2013 Oct;12(5):923-931.
(259) Chen Z, Wu J, Yang C, Fan P, Balazs L, Jiao Y, et al. DiGeorge syndrome critical region 8 (DGCR8) protein-mediated microRNA biogenesis is essential for vascular smooth muscle cell development in mice. J Biol Chem 2012 Jun 1;287(23):19018-19028.
(260) Stankiewicz TR, Schroeder EK, Kelsey NA, Bouchard RJ, Linseman DA. C-terminal binding proteins are essential pro-survival factors that undergo caspase-dependent downregulation during neuronal apoptosis. Mol Cell Neurosci 2013 Sep;56:322-332.
(261) Fernandes TG, Diogo MM, Fernandes-Platzgummer A, da Silva CL, Cabral JM. Different stages of pluripotency determine distinct patterns of proliferation, metabolism, and lineage commitment of embryonic stem cells under hypoxia. Stem Cell Res 2010 Jul;5(1):76-89.
(262) Yanes O, Clark J, Wong DM, Patti GJ, Sanchez-Ruiz A, Benton HP, et al. Metabolic oxidation regulates embryonic stem cell differentiation. Nature Chemical Biology 2010;6:411-417.
(263) Zhao FQ, Keating AF. Functional properties and genomics of glucose transporters. Curr Genomics 2007 Apr;8(2):113-128.
(264) Wong N, De Melo J, Tang D. PKM2, a Central Point of Regulation in Cancer Metabolism. Int J Cell Biol 2013;2013:242513.
(265) Yang W, Zheng Y, Xia Y, Ji H, Chen X, Guo F, et al. ERK1/2-dependent phosphorylation and nuclear translocation of PKM2 promotes the Warburg effect. Nat Cell Biol 2012 Dec;14(12):1295-1304.
(266) Kondoh H, Lleonart ME, Nakashima Y, Yokode M, Tanaka M, Bernard D, et al. A high glycolytic flux supports the proliferative potential of murine embryonic stem cells. Antioxid Redox Signal 2007 Mar;9(3):293-299.
(267) Huang S, Taylor NL, Stroher E, Fenske R, Millar AH. Succinate dehydrogenase assembly factor 2 is needed for assembly and activity of mitochondrial complex II and for normal root elongation in Arabidopsis. Plant J 2013 Feb;73(3):429-441.
(268) Berg J, Tymoczko J, Stryer L. Entry to the Citric Acid Cycle and Metabolism Through It Are Controlled. In: Freeman W, editor. Biochemistry; 2002.
(269) Vercoutere B, Durozard D, Baverel G, Martin G. Complexity of glutamine metabolism in kidney tubules from fed and fasted rats. Biochem J 2004 Mar 1;378(Pt 2):485-495.

(270) Smolkova K, Plecita-Hlavata L, Bellance N, Benard G, Rossignol R, Jezek P. Waves of gene regulation suppress and then restore oxidative phosphorylation in cancer cells. Int J Biochem Cell Biol 2011 Jul;43(7):950-968.
(271) Cairns RA, Mak TW. Oncogenic isocitrate dehydrogenase mutations: mechanisms, models, and clinical opportunities. Cancer Discov 2013 Jul;3(7):730-741.
(272) Cairns RA, Mak TW. Oncogenic isocitrate dehydrogenase mutations: mechanisms, models, and clinical opportunities. Cancer Discov 2013 Jul;3(7):730-741.
(273) Smolkova K, Jezek P. The Role of Mitochondrial NADPH-Dependent Isocitrate Dehydrogenase in Cancer Cells. Int J Cell Biol 2012;2012:273947.
(274) Mandard S, Muller M, Kersten S. Peroxisome proliferator-activated receptor alpha target genes. Cell Mol Life Sci 2004 Feb;61(4):393-416.
(275) Taguchi A, Yanagisawa K, Tanaka M, Cao K, Matsuyama Y, Goto H, et al. Identification of hypoxia-inducible factor-1 alpha as a novel target for miR-17-92 microRNA cluster. Cancer Res 2008 Jul 15;68(14):5540-5545.
(276) Shyh-Chang N, Daley GQ. Lin28: primal regulator of growth and metabolism in stem cells. Cell Stem Cell 2013 Apr 4;12(4):395-406.
(277) Dzeja P, Terzic A. Adenylate kinase and AMP signaling networks: metabolic monitoring, signal communication and body energy sensing. Int J Mol Sci 2009 Apr 17;10(4):1729-1772.
(278) Yu FX, Chai TF, He H, Hagen T, Luo Y. Thioredoxin-interacting protein (Txnip) gene expression: sensing oxidative phosphorylation status and glycolytic rate. J Biol Chem 2010 Aug 13;285(33):25822-25830.
(279) Hartl F,Ulrich, Bracher,Andreas, Hayer-Hartl,Manajit. Molecular chaperones in protein folding and proteostasis. Nature 2011;475.
(280) Kruger J, Rehmsmeier M. RNAhybrid: microRNA target prediction easy, fast and flexible. Nucleic Acids Res 2006 Jul 1;34(Web Server issue):W451-4.
(281) Martinez NJ, Gregory RI. Argonaute2 expression is post-transcriptionally coupled to microRNA abundance. RNA 2013 May;19(5):605-612.
(282) Frickel EM, Frei P, Bouvier M, Stafford WF, Helenius A, Glockshuber R, et al. ERp57 is a multifunctional thiol-disulfide oxidoreductase. J Biol Chem 2004 Apr 30;279(18):18277-18287.
(283) Laurindo FR, Pescatore LA, Fernandes Dde C. Protein disulfide isomerase in redox cell signaling and homeostasis. Free Radic Biol Med 2012 May 1;52(9):1954-1969.
(284) Sidibe A, Yin X, Tarelli E, Xiao Q, Zampetaki A, Xu Q, et al. Integrated membrane protein analysis of mature and embryonic stem cell-derived smooth muscle cells using a novel combination of CyDye/biotin labeling. Mol Cell Proteomics 2007 Oct;6(10):1788-1797.

Tables

Table 1

Summary of denaturing agents in protein solubilising buffers

Class	Common reagents	General characteristics
Chaotropes	Urea, thiourea	Disrupt hydrogen bonds and hydrophilic interactions
Reducing agents	Dithothreitol (DTT), Dithioerythritol (DTE)	Disrupt disulfide bonds between cysteine residues
Detergents		
Ionic	Sodium dodecyl sulfate (SDS)	Disrupt hydrophobic interactions
Non-ionic	Triton X-100, NP-40	Disrupt hydrophobic interactions

Table 2

Cell densities for transfections

	Per transfection	Number of cells per flask
Protein	2 T75	1.2×10^6
RNA	2 T25	4×10^5
PCA	2 T25	4×10^5

Table 3

Mastermixes for transfections

Flask size	T25	T75
Mastermix A	198ul of Transfection medium 1.5ul of mimic/control (25µM)	397ul of Transfection medium 3ul of mimic/control (25µM)
Mastermix B	200ul of Transfection medium 8ul of Lipofectamine RNAiMAX*	400ul of Transfection medium 16ul of Lipofectamine RNAiMAX*

*Lipofectamine RNAiMAX (Life Technologies, Cat No.13778-075)

Table 4

Forward and Reverse Primers

Gene primers	Forward sequence (5' to 3')	Reverse sequence (3' to 5')
Oct4	GGCGTTCTCTTTGGAAAGGTGTTC	CTCGAACCACATCCTTCTCT
Nanog	AGGGTCTGCTACTGAGATGCTCTG	CAACCACTGGTTTTTCTGCCACCG
Sox2	CACAACTCGGAGATCAGCAA	CTCCGGGAAGCGTGTACTTA
DGCR8 (specific to exon3)	GTGGATGAAGAGGCCTTGAA	TGATCACTGTCTCCGCCATA
Lats2	TCGACCTGGATGGTCATATT	ACAGCGACAGTTGGAAACAT
Rbl2	GCTCCTTACACGACGGTCTA	CATCCTCACCAAGGAATACG
SMA	GTCCCAGACATCAGGGAGTAA	TCGGATACTTCAGCGTCAGGA
SM22	GCAGTCCAAAATTGAGAAGA	CTGTTGCTGCCCATTTGAAG
Vimentin	ATGCTTCTCTGGCACGTCTT	AGCCACGCTTTCATACTGCT
FGF5	TG TCTCAGGGGATTGTAGGA	TCATCCGTAAATTTGGCA C
Isocitrate Dehydrogenase 2	AACACCGACGAGTCCATTT C	TCAGACTGCACGTCTCCATC
Serpin	CTGCGAACACTCCAAGATC A	CTGCAGCTTCTCCTTCTCGT
Cdkn1a	TCTTGCACTCTGGTGTCTGA	ATCTGCGCTTGGAGTGATAG
β-actin	CACAACTGGGACGACATGGAG	TTCATGAGGTAGTCAGTCTGG

Table 5

Taqman microRNA/gene expression assays

Gene expression assay	**Cat No.**
Aldehyde dehydrogenase 2 (Aldh2)	Mm00477463_m1
Cyclin dependent kinase inhibitor (Cdkn1a)	Mm04205640_g1
Fructose bisphosphate aldolase A (Aldoa)	Mm00833172_g1
Hypoxia-inducible factor 1-α (Hif1α)	Mm01198376_m1
Isocitrate dehydrogenase 2	Mm00612429_m1
Nanog	Mm02384862_g1
Oct4	Mm00658129_gH
Peroxisome proliferator-activated receptor α (PPARα)	Mm00440939_m1
PTEN-like mitochondrial phosphatase (PTPM1)	Mm00458631_m1
Serpin h1	Mm00438058_g1
Smooth muscle actin (SMA)	Mm01546133_m1
SM22	Mm00441661_g1
Superoxide dismutase 1 (SOD1)	Mm01344233_g1
Sox2	Mm03053810_s1
Vimentin	Mm01333430_m1
MiRNA expression assay	
mmu-miR-291a-3p	002592
mmu-miR-291b-3p	002538
mmu-miR-294-3p	001056
mmu-miR-295-3p	000189
mmu-miR-302d-3p	001939
U6	Mm01164115_g1

Table 6

Antibodies

Table	
Primary antibodies	**Company (Catalogue Number)**
Anti-Oct4 (mouse monoclonal IgG)	Santa Cruz (sc-5279)
Anti-Nanog (rabbit polyclonal IgG)	Abcam (ab80892)
Anti-Sox2 (rabbit polyclonal IgG)	Millipore (ab5603)
Anti-Serpin/hsp47 (rabbit polyclonal IgG)	Abcam (ab13519)
Anti-SOD1 (rabbit polyclonal IgG)	Santa Cruz (sc-11407)
Anti-mnSOD (rabbit polyclonal IgG)	Upstate (06-984)
Anti-Peroxiredoxin 3 (mouse monoclonal IgG)	Abcam (ab16751)
Anti-Malate dehydrogenase (rabbit polyclonal IgG)	Abcam (ab96193)
Anti-Pyruvate dehydrogenase E1α (rabbit polyclonal IgG)	Abcam (ab92696)
Anti-Protein disulfide isomerase (mouse monoclonal IgG)	Stressgen (SPA-891)
Anti-ALDOA (rabbit polyclonal IgG)	Sigma Aldrich (av48130)
Anti- Isocitrate dehydrogenase (rabbit polyclonal IgG)	Abcam (ab154932)
Anti-Tubulin (rabbit polyclonal IgG)	Abcam (ab4074)
Secondary antibodies	
Anti-rabbit IgG, HRP-linked	Dako (PO217)
Anti-mouse IgG, HRP-linked	Dako (PO260)

Table 7

DIGE analysis of wild type and DGCR8$^{-/-}$ mESCs at baseline

No	Protein name	Accession Number	Predicted Molecular Weight/pI	Observed Molecular Weight/pI	Protein identification probability	Number of unique peptides	Number of unique spectra	Number of total spectra	Fold Change	p-value	Predicted target of miR-295 cluster (miRNAwalk)
METABOLISM											
Oxidative phosphorylation											
9	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 5	NDUA5_MOUSE	13.4/7.82	15.1/7.97	100.0%	2	3	6	1.63	1.10E-02	CDS
15	Succinate dehydrogenase assembly factor 2, mitochondrial	SDHF2_MOUSE	16.6/5.35	16.2/5.26	99.0%	1	1	1	-1.52	1.60E-04	
46	Electron transfer flavoprotein subunit beta	ETFB_MOUSE	27.5/8.29	32.9/8.66	100.0%	6	7	15	1.68	3.00E-05	
53	Malate dehydrogenase, mitochondrial	MDHM_MOUSE	33.1/8.55	43.4/8.93	100.0%	16	21	43	1.63	7.60E-05	CDS
59-60	Isocitrate dehydrogenase [NADP], mitochondrial	IDHP_MOUSE	46.6/8.49	50.9/8.88	100.0%	17	20	33	2.25	1.40E-04	
90	Glycerol-3-phosphate dehydrogenase, mitochondrial	GPDM_MOUSE	76.6/5.82	73.3/5.63	100.0%	19	27	48	1.77	9.40E-06	
Glycolysis											

44	Phosphoglycerate mutase 1	PGAM1_MOUSE	28.8/6.75	33.3/7.02	100.0%	6	7	11	-1.55	1.10E-0.3	
52	Glyceraldehyde-3-phosphate dehydrogenase	G3P_MOUSE	35.8/8.45	43.8/8.93	100.0%	11	17	34	1.59	2.10E-02	
58	Fructose-bisphosphate aldolase A	ALDOA_MOUSE	39.2/8.40	52.6/8.72	100.0%	6	6	11	-1.58	2.30E-04	
63-64	Alpha-enolase	ENOA_MOUSE	47.0/6.36	57.1/6.2	100.0%	10	15	25	-1.57	1.20E-04	
Fatty Acid Associated											
49	3-hydroxyisobutyrate dehydrogenase, mitochondrial	3HIDH_MOUSE	31.7/6.01	38.4/5.78	100.0%	4	4	8	1.85	2.10E-05	
50	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	ECH1_MOUSE	32.4/6.01	39.2/6.03	100.0%	4	5	7	1.67	9.40E-06	
75	Aldehyde dehydrogenase, mitochondrial	ALDH2_MOUSE	54.3/6.05	61.3/6.00	100.0%	15	18	30	1.75	1.80E-05	
77	Aldehyde dehydrogenase X, mitochondrial	AL1B1_MOUSE	57.6/6.59	62.3/6.63	100.0%	2	2	2	2.35	2.00E-05	3'UTR
Energy i.e ATP											
45	Adenylate kinase isoenzyme 4, mitochondrial	KAD4_MOUSE	25.1/7.02	33.0/7.16	100.0%	5	6	11	1.57	6.60E-04	3' UTR
23	Nucleoside diphosphate kinase B	NDKB_MOUSE	17.2/7.17	20.0/6.57	100.0%	6	9	16	-1.87	8.20E-04	
Amino Acid											
75	Lipoamide acyltransferase	ODB2_MOUSE	46.2/5.99	61.3/6.00	100.0%	10	13	21	1.75	1.80E-05	3'UTR

	component of branched-chain alpha-keto acid dehydrogenase complex, mitochondrial										
CHAPERONE											
Mitochondrial											
5-6	10 kDa heat shock protein, mitochondrial	CH10_MOUSE	11.0/7.91	12.0/7.69	95.0%	1	1	1	-1.51	6.50E-05	
67	60kDa heat shock protein, mitochondrial	CH60_MOUSE	57.9/5.35	65.8/5.91	100.0%	2	2	2	1.57	1.30E-02	CDS
Cytosolic											
14	Prefoldin subunit 1	PFD1_MOUSE	14.1/8.14	14.6/5.69	100.0%	2	2	3	-1.57	5.30E-03	
16	Tubulin-specific chaperone A	TBCA_MOUSE	12.8/5.24	17.0/5.08	100.0%	2	2	4	-1.57	3.00E-03	
36	Cellular nucleic acid-binding protein	CNBP_MOUSE	19.6/7.76	25.0/7.96	100.0%	3	3	6	-2.08	6.80E-05	
51	DnaJ homolog subfamily C member 9	DNJC9_MOUSE	30.1/5.76	39.6/5.5	100.0%	7	10	16	-1.59	1.20E-04	
61	Serpin H1	SERPH_MOUSE	44.8/8.80	54.0/8.98	100.0%	3	4	5	-1.53	6.20E-06	
70	Peptidyl-prolyl cis-trans isomerase FKBP4	FKBP4_MOUSE	51.6/5.54	62.7/5.52	100.0%	4	4	6	2.12	1.90E-05	CDS
71-72/74	Protein disulfide-isomerase A3	PDIA3_MOUSE	54.2/5.69	66.5/5.66	100.0%	26	31	54	1.73	1.30E-05	
73	DnaJ homolog subfamily C member 7	DNJC7_MOUSE	56.3/6.12	66.5/5.74	100.0%	13	15	20	1.96	6.80E-05	
80	Calreticulin	CALR_MOUSE	46.3/4.33	72.8/4.40	100.0%	11	11	17	1.96	1.70E-03	
84-	78 kDa glucose-	GRP78_M	70.4/5.01	73.3/5.3	100.0%	32	51	88	1.74	2.20E-05	

85/87	regulated protein	OUSE									
86-87	Protein disulfide-isomerase A4	PDIA4_MOUSE	72.0/5.06	73.6/5.18	100.0%	28	33	56	1.55	2.30E-04	
88-89	Heat shock cognate 71kDa protein	HSP7C_MOUSE	70.7/5.37	72.8/5.58	100.0%	8	8	11	1.74	1.1E-02	
96	Hypoxia up-regulated protein 1	HYOU1_MOUSE	107.6/5.06	111.2/5.12	100.0%	31	40	63	1.53	3.30E-02	
Aid Chaperone function											
7	Ubiquitin-like protein	UBL5_MOUSE	8.5/8.58	11.5/8.98	95.0%	1	1	1	1.99	1.30E-05	
18-19	Peptidyl-prolyl cis-trans isomerase-like 3	PPIL3_MOUSE	18.1/6.29	18.2/6.34	100.0%	2	2	4	2.36	1.00E-03	
20-21	Peptidyl-prolyl cis-trans isomerase A	PPIA_MOUSE	18.0/7.88	18.3/7.96	100.0%	7	9	19	-1.58	2.20E-03	
ANTIOXIDANT											
37	Peroxiredoxin-1	PRDX1_MOUSE	22.2/8.26	26.8/8.65	100.0%	12	13	24	2.08	2.10E-02	
48	Carbonyl reductase [NADPH] 3	CBR3_MOUSE	31.0/6.15	36.5/5.52	100.0%	4	4	5	1.65	6.80E-05	
17	Superoxide dismutase [Cu-Zn]	SODC_MOUSE	15.8/6.03	17.8/5.7	100.0%	4	6	12	-1.75	7.80E-04	
38	Superoxide dismutase [Mn], mitochondrial	SODM_MOUSE	22.2/7.30	27.1/7.94	95.0%	5	5	9	1.69	2.10E-05	3'UTR
ASSOCIATED WITH TRANSCRIPTION/TRANSLATION											
DNA Processing											
10-12	Histone H4	H4_MOUSE	11.7/8.63	13.9/6.92	100.0%	3	3	5	1.5	3.90E-03	
28	Mediator of RNA polymerase II transcription subunit 22	MED22_MOUSE	22.3/4.56	22.3/4.56	100.0%	3	3	5	-1.63	1.80E-02	

34	Chromobox protein homolog 3	CBX3_MOUSE	20.9/5.13	24.2/5.06	100.0%	7	8	14	-1.65	6.20E-06	
76	RuvB-like 1	RUVB1_MOUSE	50.2/6.02	61.0/6.43	100.0%	11	14	22	1.55	1.00E-04	
91	Far upstream element-binding protein 1	FUBP1_MOUSE	68.4/7.85	75.1/7.66	100.0%	12	15	23	-1.52	1.10E-04	
94	Ribonucleoside-diphosphate reductase large subunit	RIR1_MOUSE	90.2/6.27	90.2/6.27	99.0%	1	1	1	1.52	1.30E-02	5'UTR/ CDS
RNA Processing											
4	U6 snRNA-associates Sm-like protein LSM2	LSM2_MOUSE	10.8/6.05	10.8/6.05	100.0%	2	2	3	-1.59	2.30E-04	CDS
24	Transcription factor BTF3 homolog 4	BT3L4_MOUSE	17.3/5.95	21.6/5.52	100.0%	3	3	5	-1.8	1.60E-02	
76	Cleavage stimulation factor subunit 1	CSTF1_MOUSE	48.3/6.12	61.0/6.43	100.0%	8	10	18	1.55	1.00E-04	
92/ 93	ATP-dependent RNA helicase DDX3X	DDX3X_MOUSE	73.0/6.73	77.5/7.3	100.0%	29	40	71	1.84	1.70E-04	
PROTEIN SYNTHESIS											
Mitochondrial											
39	Peptidyl-tRNA hydrolase 2 mitochondrial	PTH2_MOUSE	19.5/6.95	27.2/7.43	100.0%	3	5	7	1.51	1.50E-03	
47	3-hydroxyacyl-CoA dehydrogenase type-2	HCD2_MOUSE	27.3/8.56	31.0/8.5	100.0%	6	8	16	2.03	1.20E-04	CDS
Cytosolic											
8	40S ribosomal protein S21	RS21_MOUSE	9.1/8.71	12.8/8.91	100.0%	2	2	3	1.87	5.90E-03	
26-27	Eukaryotic translation initiation factor 5A-1	IF5A1_MOUSE	16.8/5.08	19.2/5.14	100.0%	9	16	26	-2.11	9.50E-06	
31	Eukaryotic translation initiation factor 1A, X-	IF1AX_MOUSE	16.5/5.07	21.5/5.13	100.0%	5	6	11	-1.83	3.90E-06	

	chromosomal										
34	M-phase phosphoprotein 6	MPH6_MOUSE	19.1/5.25	24.2/5.06	100.0%	4	4	8	-1.65	6.20E-06	
35	39S ribosomal protein L12, mitochondrial	RM12_MOUSE	21.7/9.34	23.8/5.57	90.0%	1	1	1	2.03	6.70E-04	
75	Myc-induced nuclear antigen	MINA_MOUSE	53.5/5.96	61.3/6.00	100.0%	10	12	19	1.75	1.80E-05	
95	U3 small nucleolar ribonucleoprotein protein MPP10	MPP10_MOUSE	78.7/4.78	89.0/4.78	100.0%	14	19	28	1.78	9.20E-03	
CELL CYCLE ASSOCIATED											
7	Small nuclear ribonucleoprotein G	RUXG_MOUSE	8.5/8.98	11.5/8.98	95.0%	1	1	2	1.99	1.30E-05	
13	CDKN2AIP N-terminal-like protein	C2AIL_MOUSE	13.2/5.14	12.4/5.13	100.0%	2	2	4	-1.55	1.90E-02	3'UTR
29-30	Calmodulin	CALM_MOUSE	16.8/4.09	20.3/4.03	100.0%	2	4	5	1.61	4.10E-02	3'UTR
40-41	Translationally-controlled tumour protein	TCTP_MOUSE	19.5/4.76	29.0/4.89	100.0%	4	9	17	-1.74	6.20E-06	
42	GTP-binding nuclear protein Ran	RANG_MOUSE	24.0/5.15	33.1/5.21	100.0%	2	2	3	-1.91	9.40E-06	
66	Transcription factor E2F4	E2F4_MOUSE	43.8/4.65	66.3/4.50	100.0%	2	2	3	1.69	2.20E-03	
83	Lamin-B1	LMNB1_MOUSE	66.3/5.11	72.5/5.3	100.0%	25	32	50	1.71	7.30E-04	
POST-TRANSLATIONAL MODIFICATIONS											
19	SUMO-conjugating enzyme UBC9	UBC9_MOUSE	18.0/8.87	18.9/8.58	100.0%	3	6	11	-1.53	8.80E-06	
ACTIN ASSOCIATED											
22	Destrin	DEST_MO	18.4/8.19	19.7/7.43	100.0%	4	5	6	-1.54	5.40E-04	

		USE									
55-57	Actin, cytoplasmic 1	ACTB_MOUSE	41.7/5.29	49.3/5.38	100.0%	10	18	36	1.64	1.10E-4	
78	Adenylyl cyclase-associated protein 1	CAP1_MOUSE	51.4/7.30	64.0/7.62	100.0%	6	7	10	0.89	1.90E-05	
SMOOTH MUSCLE ASSOCIATED											
54	Calponin-3	CNN3_MOUSE	36.4/5.46	48.5/5.2	100.0%	10	12	25	1.5	2.20E-02	
81-82	Vimentin	VIME_MOUSE	53.6/5.05	69.7/5.3	100.0%	26	38	60	2.08	7.20E-05	CDS
IMMUNE											
65	CD2 antigen cytoplasmic tail-binding protein 2	CD2B2_MOUSE	37.7/4.48	67.1/4.5	100.0%	5	5	7	1.65	3.10E-03	3' UTR
CALCIUM SIGNALLING											
2	Protein S100-A6	S10A6_MOUSE	10.1/5.3	10.1/5.3	95.0%	1	1	2	1.89	1.30E-02	
33	Calcineurin B homologous protein 1	CHP1_MOUSE	22.4/4.97	25.1/4.78	100.0%	7	10	14	-1.88	5.60E-05	
OTHER											
9	Hemoglobin subunit alpha	HBA_MOUSE	15.0/8.08	15.1/7.97	100.0%	3	4	7	1.63	1.10E-02	
25	Ferritin light chain 1	FRIL1_MOUSE	20.8/5.65	18.0/5.3	100.0%	4	5	6	-1.59	1.50E-03	
32	Prostaglandin E synthase 3	TEBP_MOUSE	18.7/4.33	24.5/4.30	100.0%	4	5	6	-1.57	1.10E-04	
43	Proteasome subunit alpha type-6	PSA6_MOUSE	27.3/6.34	27.3/6.34	100.0%	6	7	12	-1.5	9.80E-05	
62	Annexin A7	ANXA7_MOUSE	49.9/5.91	56.0/5.53	100.0%	4	6	10	2.22	6.80E-05	
68-69	Tubulin alpha-1B chain	TBA1B_MOUSE	50.1/4.94	64.6/5.27	100.0%	14	19	33	1.56	7.90E-05	

Table 8

Average percentages (± standard deviation) of metabolites produced by Wild type and DGCR8$^{-/-}$ mESCs over 48 hours.

	Average percentages ± standard deviation		
	Wild Type	DGCR8$^{-/-}$	p-value (t-test)
Leucine	2%± 0.74	2%± 0.59	0.50
Valine	2% ± 0.76	1%±0.53	0.52
Iso-leucine	2% ± 0.69	2%± 0.43	0.96
Lactate	-51% ± 7.85	-48%± 9.92	0.26
Alanine	-3% ± 0.62	-5%± 1.83	0.04
Acetate	-3% ± 1.32	-3%± 0.80	0.81
Glutamine	7% ± 1.74	9%± 3.09	0.65
Choline	3% ± 0.56	2%± 0.26	0.22
Glucose	27% ± 6.84	26%± 7.73	0.58
Histidine	0% ± 0.02	0%± 0.22	0.94
Tyrosine	1% ± 0.24	1%± 0.30	0.83
Phenylalanine	1% ± 0.30	1%± 0.62	0.58

Negative values indicate that metabolites are excreted, whereas positive values indicate that metabolites are taken up by the cells. Values have been normalised to cell numbers after 48 hours of growth. Metabolites with a significant difference between the two cell types are highlighted in red. p-values were calculated using an independent student's t-text, a value of less than 0.05 was taken as significant

Table 9

DIGE analysis of DGCR8$^{-/-}$ mESCs transfected with ESCC miRNAs

No	Protein name	Accession Number	Predicted Molecular Weight/pI	Observed Molecular Weight/pI	Protein identification probability	Number of unique peptides	Number of unique spectra	Number of total spectra	Fold change	ANOVA p-value	No. of ESCC miRNAs	ANOVA p-value
METABOLISM												
Oxidative phosphorylation												
1	Cytochrome c oxidase subunit 5A, mitochondrial	COX5A_MOUSE	16.1/5.01	12.3/4.96	22.60%	4	4	6	1.08	5.00E-01		
11	ATP synthase subunit d, mitochondrial	ATP5H_MOUSE	18.8/5.53	21.3/5.49	32.90%	5	7	12	1.16	4.60E-02	1	**
16	ATP synthase subunit b, mitochondrial	AT5F1_MOUSE	28.9/8.55	25.3/8.63	19.50%	5	6	11	1.15	1.80E-02		
17/18	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial	NDUS3_MOUSE	30.1/5.45	27.2/5.47	49.30%	15	22	40	1.13	1.70E-01	2	**
20	Electron transfer flavoprotein subunit beta	ETFB_MOUSE	27.6/8.29	28.0/8.72	60.00%	16	24	42	1.47	2.20E-04	4	*
23	ATP synthase subunit gamma, mitochondrial	ATPG_MOUSE	32.9/8.87	31.6/8.84	28.20%	7	9	14	1.21	1.70E-02		
28	Malate dehydrogenase, mitochondrial	MDHM_MOUSE	35.6/8.55	37.4/8.36	57.40%	17	23	50	1.36	2.50E-02	1	*
39	ATP synthase subunit beta, mitochondrial	ATPB_MOUSE	56.3/4.99	50.7/4.91	40.60%	14	20	33	1.07	3.10E-01		
42	Propionyl-CoA	PCCB_M	58.4/6.53	54.3/7.5	20.20%	9	10	17	1.36	2.70E-		

	carboxylase beta chain, mitochondrial	OUSE		4						02		
Glycolysis												
29	Glyceraldehyde-3-phosphate dehydrogenase	G3P_MOUSE	35.8/8.45	38.1/8.79	33.30%	12	20	31	1.4	1.40E-03	4	*
33	Fructose-bisphosphate aldolase A	ALDOA_MOUSE	39.3/8.4	42.8/8.68	73.90%	26	39	65	-1.47	5.90E-03	1	*
36	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	ODPA_MOUSE	43.2/6.78	46.3/8.87	22.60%	8	11	20	-1.11	3.10E-01		
45	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	ODP2_MOUSE	67.9/5.7	59.3/5.49	42.60%	24	29	45	-1.17	2.60E-01		
Fatty Acid Associated												
23	2,4-dienoyl-CoA reductase, mitochondrial	DECR_MOUSE	36.2/8.78	31.5/7.56	31.50%	8	11	18	1.21	1.70E-02	1	*
23	D-beta-hydroxybutyrate dehydrogenase, mitochondrial	BDH_MOUSE	38.3/8.62	31.5/7.56	32.20%	8	9	12	1.21	1.70E-02	1	*
26	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	ECH1_MOUSE	36.1/6.01	32.3/5.65	25.30%	7	8	11	1.41	5.00E-03	1	*
50	Peroxisomal targeting signal 1 receptor	PEX5_MOUSE	70.8/4.44	70.8/4.44	8.14%	4	4	6	1.18	3.00E-02	1	**
Energy i.e ATP metabolism												
19	GTP:AMP	KAD4_	25.1/7.02	27.3/8.39	15.20%	4	6	8	1.21	6.20E-02	2	**

	phosphotransferase AK4, mitochondrial	MOUSE										
36	Multifunctional protein ADE2	PUR6_MOUSE	47.0/7.09	46.3/8.87	25.60%	11	13	22	-1.11	3.10E-01		
CHAPERONE												
4/5	Peptidyl-prolyl cis-trans isomerase A	PPIA_MOUSE	18.0/7.73	18.0/7.73	66.50%	18	30	56	-1.12	2.70E-01		
8	60S ribosome subunit biogenesis protein NIP7 homolog	NIP7_MOUSE	20.5/8.35	19.7/8.66	57.20%	7	11	20	-1.1	2.20E-01		
25	DnaJ homolog subfamily C member 9	DNJC9_MOUSE	30.1/5.76	32.0/5.41	33.80%	10	15	25	1.5	1.70E-04	4	**
37	DnaJ homolog subfamily A member 1	DNJA1_MOUSE	44.9/6.65	48.2/8.78	39.70%	16	30	63	-1.02	8.6E-01	1	*
44	Protein disulfide-isomerase	PDIA1_MOUSE	57.1/4.72	57.5/4.85	10.40%	7	7	10	-1.05	8.6E-01		
52	Heat shock cognate 71 kDa protein	HSP7C_MOUSE	70.9/5.37	68.5/5.36	41.30%	30	46	75	-1.22	5.80E-01		
53	Stress-70 protein, mitochondrial	GRP75_MOUSE	73.5/5.44	68.7/5.39	10.30%	8	8	13	-1.24	1.20E-02	2	*
CELLULAR TRANSPORT												
Mitochondrial Transport												
7	Mitochondrial import receptor subunit TOM22 homolog	TOM22_MOUSE	15.5/4.29	18.1/5.00	16.20%	2	3	6	1.01	8.00E-01		
23/27	Voltage-dependent anion-selective channel protein 1	VDAC1_MOUSE	32.4/8.55	32.3/8.55	13%/74%	24	33	73	1.23	2.30E-03	3	*
Protein transport												
48	Vesicle-fusing ATPase	NSF_MOUSE	82.6/6.52	31.6/8.84	26.90%	17	20	31	1.04	7.20E-01		

REDOX SIGNALLING												
9/10/13	Peroxiredoxin-1	PRDX1_MOUSE	22.2/8.26	23.0/8.58	71.90%	45	58	119	1.18	2.50E-01	3	*
12	Peroxiredoxin-2	PRDX2_MOUSE	21.8/5.2	21.9/4.98	31.80%	6	8	19	1.19	1.20E-01	1	*
15	Thioredoxin-dependent peroxide reductase, mitochondrial	PRDX3_MOUSE	28.1/5.73	24.1/5.4	28.80%	8	12	22	1.27	1.10E-01	1	***
22	Protein SCO1 homolog, mitochondrial	SCO1_MOUSE	31.7/9.7	29.2/7.61	13.00%	4	5	9	1.15	8.10E-01		
ASSOCIATED WITH TRANSCRIPTION/TRANSLATION												
DNA processing												
2	Histone H2B type 1-P	H2B1B_MOUSE	13.9/10.32	13.9/5.06	38.10%	4	4	5	-1.01	9.10E-01		
3	Histone H4	H4_MOUSE	11.4/11.36	14.5/5.00	29.10%	3	3	5	1.09	4.00E-01		
41	Pre-mRNA-processing factor 19	PRP19_MOUSE	55.2/6.14	54.3/6.37	18.50%	9	11	20	1.13	2.20E-01	2	*
RNA processing												
36	Heterogeneous nuclear ribonucleoprotein D0	HNRPD_MOUSE	38.4/7.61	46.3/8.87	42.30%	14	14	23	-1.11	3.10E-01		
47	Probable ATP-dependent RNA helicase DDX5	DDX5_MOUSE	69.3/9.06	61.8/8.83	25.20%	17	20	32	1.1	3.40E-01	1	*
49	Heterogeneous nuclear ribonucleoprotein M	HNRPM_MOUSE	77.7/8.81	64.7/8.63	22.80%	15	20	33	-1.17	1.70E-01		
51	PC4 and SFRS1-interacting protein	PSIP1_MOUSE	59.7/9.15	68.2/8.69	13.90%	6	6	9	-1.2	1.20E-02		

51	Splicing factor 1	SF01_MOUSE	70.4/9.14	68.2/8.69	6.58%	5	5	7	-1.2	1.20E-02		
55/56	Staphylococcal nuclease domain-containing protein 1	SND1_MOUSE	102.1/7.14	74.8/8.87	24.30%	35	39	59	-1.1	4.90E-01	2	*
58	Transcription intermediary factor 1-beta	TIF1B_MOUSE	88.8/5.52	79.2/5.33	19.90%	14	14	25	-1.13	1.80E-01		
Protein synthesis												
6	Eukaryotic translation initiation factor 5A-1	IF5A1_MOUSE	16.8/5.08	16.1/4.94	35.70%	8	11	19	-1.39	8.50E-04	3	*
35	Methionine aminopeptidase 1	MAP11_MOUSE	43.2/6.8	45.3/8.61	24.60%	7	9	17	1.02	8.30E-01		
38	Nuclease-sensitive element-binding protein 1	YBOX1_MOUSE	35.7/9.87	50.6/8.85	43.20%	9	12	19	1.13	2.90E-01		
54	C-1-tetrahydrofolate synthase, cytoplasmic	C1TC_MOUSE	101.2/6.73	72.9/8.75	38.00%	23	31	49	-1.05	5.70E-01		
CELL CYCLE												
7	S-phase kinase-associated protein 1	SKP1_MOUSE	18.7/4.4	18.1/5.00	7.98%	2	3	4	1.01	8.00E-01		
51	Lamina-associated polypeptide 2, isoforms alpha/zeta	LAP2A_MOUSE	75.2/8.35	68.2/8.69	16.90%	8	9	13	-1.2	1.20E-02		
54	Kinesin-like protein KIF2C	KIF2C_MOUSE	101.2/6.73	72.9/8.75	14.00%	9	9	13	-1.05	5.70E-01	1	*
MEIOSIS												
16	Meiotic nuclear division protein 1 homolog	MND1_MOUSE	23.9/8.37	25.3/8.63	27.80%	6	6	10	1.15	1.80E-02		
APOPTOSIS												
46	Apoptosis-inducing	AIFM1_	66.8/7.27	61.7/8.7	36.80%	20	28	44	1.13	1.60E-	1	**

	factor 1, mitochondrial	MOUSE		7						01		
POST-TRANSLATIONAL MODIFIER												
30	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit	PP1A_MOUSE	37.5/5.94	39.5/5.4	26.10%	10	10	16	1.08	3.10E-01	1	*
CYTOSKELETAL												
43/48	Moesin	MOES_MOUSE	67.8/6.24	57.3/5.92	24.70%	29	33	57	-1.21	6.90E-02		
57	LIM domain and actin-binding protein 1	LIMA1_MOUSE	84.1/6.18	84.1/6.18	21.20%	16	19	30	-1.03	8.90E-01		
SMOOTH MUSCLE ASSOCIATED												
31,32	Calponin-3	CNN3_MOUSE	36.4/5.46	36.4/5.46	39.10%	19	31	51	1.23	2.30E-01		
OTHER												
14	Membrane-associated progesterone receptor component 1	PGRC1_MOUSE	21.7/4.54	24.1/4.53	51.30%	14	15	25	1.23	3.90E-02	3	*
21	Protein PBDC1	PBDC1_MOUSE	22.2/4.49	22.2/4.49	31.30%	4	4	7	1.02	8.00E-01		
24	Sulfatase-modifying factor 2	SUMF2_MOUSE	34.7/6.27	32.3/5.39	26.30%	9	12	19	1.16	5.20E-02	2	*
34	Tropomyosin alpha-1 chain	TPM1_MOUSE	32.7/4.69	43.1/4.64	52.80%	19	21	33	1.07	1.80E-01	3	*
40	V-type proton ATPase subunit B, brain isoform	VATB2_MOUSE	56.6/5.57	53.0/5.48	14.50%	8	9	16	1.22	1.60E-01	1	*
42	Cytosol aminopeptidase	AMPL_MOUSE	56.1/7.61	54.3/7.54	22.90%	10	11	20	1.36	2.70E-02		

* denotes p<0.05, ** denotes p<0.01, *** denotes p<0.001 p-values were calculated using an one-way ANOVA

Table 10

Efficiency of transfections based on average Ct values from qPCR(n=6)

		Average Ct values		
		Wild type mESCs	**DGCR8$^{-/-}$ mESCs**	**Transfected DGCR8$^{-/-}$ mESCs**
	Plate			
miR-291a	**1**	21.805	33.929	17.703
	2	21.762	30.791	17.312
	3	22.201	24.828	16.261
miR-291b	**1**	28.492	33.943	21.333
	2	28.362	34.368	22.696
	3	28.269	36.144	21.671
miR-294	**1**	24.138	34.938	19.497
	2	24.049	35.338	21.320
	3	24.227	37.641	19.625
miR-295	**1**	21.272	32.580	17.272
	2	21.335	31.912	17.995
	3	21.228	31.783	17.572
miR-302d	**1**	39.425	Undetermined	27.961
	2	Undetermined	Undetermined	29.607
	3	38.465	Undetermined	28.227
U6	**1**	20.544	19.137	
	2	20.389	19.975	
	3	20.477	20.173	

Table 11

Average cell numbers of the DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs

	Final cell number (x 10^6)	Average cell number (x 10^6)	Standard deviation (x 10^6)	Percentage of control
Untransfected	4.65	4.24	0.34	117.84
	4.25			
	4.23			
	3.815			
Control	5.375	3.60	1.38	100.00
	3.55			
	3.44			
	2.015			
miR-291a	4.95	3.83	0.79	106.47
	3.18			
	3.815			
	3.365			
miR-291b	3.485	3.39	0.34	94.23
	3.8			
	3.25			
	3.015			
miR-294	4.215	3.54	0.54	98.47
	3.715			
	3.015			
	3.215			
miR-295	4.7	3.63	1.02	101.01
	4.15			
	2.375			
	3.3			
miR-302d	3.73	3.52	0.41	97.91
	4			
	3.135			
	3.215			

ANOVA =0.787

Table 12

Secretome analysis of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs

		Untransfected		Control		miR-291a		miR-291b		miR-294		miR-295		miR-302d		ANOVA
			s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	
Taken up	Choline	0.18	0.06	0.13	0.08	0.21	0.05	0.26	0.08	0.23	0.07	0.25	0.11	0.26	0.07	0.272
	Glucose	1.41	1.22	1.19	1.00	1.44	1.36	1.48	1.58	1.87	0.98	1.19	0.62	2.09	2.05	0.957
	Glutamine	1.23	0.39	1.41	0.07	1.03	0.69	1.24	0.75	1.29	0.49	0.94	0.46	1.31	1.00	0.951
	Histidine	0.04	0.02	0.05	0.02	0.04	0.05	0.05	0.04	0.06	0.02	0.03	0.02	0.05	0.06	0.942
	Iso-Leucine	0.16	0.10	0.16	0.07	0.14	0.14	0.14	0.14	0.17	0.06	0.11	0.06	0.23	0.21	0.899
	Leucine	0.22	0.09	0.22	0.09	0.22	0.16	0.23	0.17	0.26	0.10	0.20	0.11	0.30	0.23	0.97
	Phenylalanine	0.07	0.05	0.04	0.06	0.05	0.08	0.05	0.08	0.06	0.04	0.04	0.04	0.09	0.11	0.953
	Pyruvate	0.34	0.05	0.38	0.14	0.42	0.12	0.46	0.08	0.46	0.10	0.43	0.14	0.46	0.06	0.585
	Tyrosine	0.06	0.05	0.06	0.04	0.05	0.08	0.05	0.07	0.07	0.03	0.05	0.03	0.09	0.09	0.962
	Valine	0.18	0.11	0.18	0.09	0.16	0.16	0.16	0.16	0.21	0.07	0.14	0.07	0.26	0.24	0.923
Secreted	Acetate	0.30	0.09	0.26	0.01	0.32	0.05	0.36	0.18	0.32	0.11	0.32	0.15	0.32	0.09	0.958
	Alanine	0.53	0.17	0.53	0.10	0.28	0.03	0.58	0.27	0.29	0.10	0.32	0.17	0.26	0.07	0.02
	Fumarate	0.0015	0.0009	0.0029	0.0009	0.0036	0.0011	0.0042	0.0036	0.0031	0.0022	0.0035	0.0014	0.0035	0.0020	0.633
	Lactate	6.20	1.69	5.90	1.12	8.94	1.95	10.14	1.79	9.74	1.82	9.88	1.80	8.53	1.31	0.01
	Threonine	0.08	0.14	0.09	0.24	0.27	0.29	0.57	0.28	0.37	0.24	0.49	0.16	0.29	0.26	0.072

s.d represents standard deviation

Table 13

Intracellular metabolite analysis of DGCR8$^{-/-}$ mESCs transfected with the ESCC miRNAs

	Control		miR-294		miR-302d		ANOVA
	Mean	s.d.	Mean	s.d	Mean	s.d.	
Acetate	0.89	0.55	1.36	1.02	1.22	0.89	0.73
ADP+ATP	1.41	0.34	2.37	0.81	2.07	0.57	0.13
Alanine	3.70	1.13	2.72	1.17	2.01	0.53	0.01
Asparate	0.31	0.16	0.24	0.09	0.25	0.20	0.80
Choline	0.08	0.02	0.09	0.03	0.11	0.10	0.79
Creatine	0.83	0.07	1.27	0.41	1.29	0.31	0.10
Formate	0.72	0.50	1.06	0.53	0.84	0.37	0.60
Fumarate	0.05	0.01	0.11	0.07	0.08	0.04	0.25
Glucose	5.65	2.31	6.75	2.83	5.03	2.51	0.64
Glutamate	5.17	0.57	8.60	2.65	6.94	0.88	0.05
Glutamine	4.48	1.01	5.23	1.66	4.41	0.72	0.58
Glycine	4.19	2.41	2.24	0.65	1.86	0.65	0.11
Glyerophosphocholine	0.35	0.23	1.04	0.69	0.73	0.22	0.14
Lactate	7.56	2.61	19.87	5.92	17.78	6.01	0.02
NAD(H)	0.14	0.02	0.28	0.13	0.36	0.12	0.04
Phosphocholine	2.93	0.24	4.18	1.23	4.06	1.52	0.28
Phosphocreatine	0.77	0.17	1.25	0.70	1.06	0.08	0.31
Succinate	0.23	0.04	0.45	0.15	0.36	0.11	0.06

s.d represents standard deviation

Table 14

Proteins identified in both DIGE studies

Protein name	Accession Number	DGCR8$^{-/-}$ vs Wild Type	Targeted by miRNAs based on proteomics	Predicted target of...*
METABOLISM				
Oxidative phosphorylation				
Electron transfer flavoprotein subunit beta	ETFB_MOUSE	1.68	miR-291b, 294, 295, 302d	
Malate dehydrogenase, mitochondrial	MDHM_MOUSE	1.63	miR-291a, 291b,302d	miR-291a[4,5], 291b[4]
Glycolysis				
Glyceraldehyde-3-phosphate dehydrogenase	G3P_MOUSE	1.59	miR-291a, 291b, 294, 295, 302d	miR-302d[5]
Fructose-bisphosphate aldolase A	ALDOA_MOUSE	1.58	miR-295	
Fatty Acid Associated				
Delta (3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	ECH1_MOUSE	1.67	miR-295	
CHAPERONE				
Peptidyl-prolyl cis-trans isomerase A	PPIA_MOUSE	-1.58		
Heat shock cognate 71kDa protein	HSP7C_MOUSE	1.74		
REDOX SIGNALLING				
Peroxiredoxin-1	PRDX1_MOUSE	2.08	miR-291a, 291b, 295, 302d	
ASSOCIATED WITH TRANSCRIPTION/TRANSLATION				
DNA processing				
Histone H4	H4_MOUSE	1.5		miR-291b[1]
Protein synthesis				

Eukaryotic translation initiation factor 5A-1	IF5A1_MOUSE	-2.11	miR-291a, 291b, 302d	
SMOOTH MUSCLE ASSOCIATED				
Calponin-3	CNN3_MOUSE	1.5		miR-291b[4,5],294[4,5], 295[4,5]
OTHER				
Cytosol aminopeptidase	AMPL_MOUSE	1.62		miR-291a[4],291b[4],294[4,] 295[4], 302d[4]

Information taken from five different microRNA databases, Targetscan[1], MicroRNA.org[2], DIANA[3], mirwalk[4] and mirnamap[5]. Dec, 2013